

1 GGATCCCAGGGAACGTGACC ATG GTC GTA GGG ATG ACT TGA CAGTTTCAACGGGGTGCACCACCGTTGCGC 72
 1 M V V G M T * 7
 73 TCAGAAGGCATACGTTGGTGAACACGTCGGAAAGCTGGGAGGTGAATCTG ATG GCT GGC GAC CAA GAG CTG 144
 1 M A G D Q E L 7
 145 GAA CTG CGG TTC GAC GTT CCT CTT TAC ACG CTT GCC GAG GCA TCG CGG TAC CTG GTG GTT 204
 8 E L R F D V P L Y T L A E A S R Y L V V 27
 205 CCC CGC GCC ACC CTG GCT ACG TGG GCT GAC GGC TAC GAG CGT CGG CCG GCC AAC GCA CCG 264
 28 P R A T L A T W A D G Y E R R P A N A P 47
 265 GCG GTC CAG GGG CAA CCG ATC GCC TTT GAC GCC TAT TCG GTC GCG CAG CTT TTT GGC GAC 324
 48 A V Q G Q P I A F D A Y S V A Q L F G D 67
 325 GTC ACT GGT GCC CGC GTT GCG GGC GTC CAG CCG CAG CGA CAC CAC ATA CGG CCG GTC CCG 384
 68 V T G A R V A G V Q P Q R H H I R P V R 87
 385 TTG CGG GGG CCG TTG GGT GGG GTT GGG TGC CTC CGT CAC CCC AGG CAG TTC GCT GGC TAT 444
 88 L R G P L G G V G C L R H P R Q F A G Y 107
 445 TTG TCG CAG TAG CGCGACGGCATTGTGTCG ATG TCT TGG TAG CTAGCATCCGGTCGGGGGGCCGCTACCAGCG 515
 108 L S Q * M S W * 4
 516 CCAGCGCCGGGGCTCCCCGGTCCGGGTAGTGC GCGTCGAGTTGGTCGTGGACCAGCA ATG ACT GCG ACC CCG 587
 1 M T A T R 5
 588 CGA CTT CGA AAC CGC CAC CGG TTA GAT TCC CCG ACT GCG TCA TCG CCA GGT AAA CCG CCG 647
 6 R L R N R H R L D S P T A S S P G K P P 25
 648 GCA CTA ACG CCA GCA ACC AAC CCG TGA AGACCAACCAACGGCACCTGCGCAGGTTGCGGCTCAACCGCATC 718
 26 A L T P A T N P * 34
 719 ATG AAC TGC TGG ATT TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT GCC CGC GAG CCT ACC 778
 1 M N C W I S D S P Y S R A V R A R E P T 20
 779 GAA GAT CGC GTG CAT GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA GTT GGC GGC GCC GAG 838
 21 E D R V H A F G V D R T A P G V G G A E 40
 839 GGC CGA GAT GGC AGG ATG ACG GAT CGT CGG GGG CGG GAA CTC CCA GGC CGC CGG ACC GTC 898
 41 G R D G R M T D R R G R E L P G R R T V 60
 899 GCA AAC CCG TCG CAA ACC CGT CGC AAA CCG TAA GGAGTCATCC ATG AAG ACA GGC ACC GCG 959
 61 A N P S Q T R R K P * M K T G T A 6
 960 ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG GGG GCC GCC GTT 1019
 7 T T R R R L L A V L I A L A L P G A A V 26
 1020 GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG TGC GCG GCC AGC GAA GTG 1079
 27 A L L A E P S A T G A S D P C A A S E V 46
 1080 GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT TCA CAC CCA GAG 1139
 47 A R T V G S V A K S M G D Y L D S H P E 66
 1140 ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG TCG GTC GCA TCG 1199
 67 T N Q V M T A V L Q Q Q V G P G S V A S 86
 1200 CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT CC 1243
 87 L K A H F E A N P K V A S D 100

SEQ ID No.1

FIGURE 1

REPLACEMENT SHEET (RULE 26)

Insert of the clone containing DP428 and contained in seq1

1/1	31/11
GAT CGC CTT TGA CGC CTA TTC GGT CGC GCA	GCT TTT TGG CGA CGT CAC TGG TGC CCG CGT
asp arg leu OPA arg leu phe gly arg ala	ala phe trp arg arg his trp cys pro arg
61/21	91/31
TGC GGG CGT CCA GCC GCA GCG ACA CCA CAT	ACG GCC GGT CCG GTT GCG GGG GCC GTT GGG
cys gly arg pro ala ala ala thr pro his	thr ala gly pro val ala gly ala val gly
121/41	151/51
TGG GGT TGG GTG CCT CCG TCA CCC CAG GCA	GTT CGC TGG CTA TTT GTC GCA GTA GCG CGA
trp gly trp val pro pro ser pro gln ala	val arg trp leu phe val ala val ala arg
181/61	211/71
CGG CAT TGT CGA TGT CTT GGT AGC TAG CAT	CCG GTC GGG GGG CCG CTA CCA GCG CCA GCG
arg his cys arg cys leu gly ser AMB his	pro val gly gly pro leu pro ala pro ala
241/81	271/91
CCG GGG CTC CCC GGT CCG GGT AGT GCG CGT	CGA GTT GGT CGT GGA CCA GCA ATG ACT GCG
pro gly leu pro gly pro gly ser ala arg	arg val gly arg gly pro ala met thr ala
301/101	331/111
ACC CGG CGA CTT CGA AAC CGC CAC CGG TTA	GAT TCC CCG ACT GCG TCA TCG CCA GGT AAA
thr arg arg leu arg asn arg his arg leu	asp ser pro thr ala ser ser pro gly lys
361/121	391/131
CCG CCG GCA CTA ACG CCA GCA ACC AAC CCG	TGA AGA CCA ACC AAC GGC ACC TGC GCA GGT
pro pro ala leu thr pro ala thr asn pro	OPA arg pro thr asn gly thr cys ala gly
421/141	451/151
TGC GGC TCA ACC GCA TCA TGA ACT GCT GGA	TTT CGG ACT CCC CGT ACT CTC GCG CAG TGC
cys gly ser thr ala ser OPA thr ala gly	phe arg thr pro arg thr leu ala gln cys
481/161	511/171
GTG CCC GCG AGC CTA CCG AAG ATC GCG TGC	ATG CGT TCG GCG TGG ACC GCA CAG CAC CTG
val pro ala ser leu pro lys ile ala cys	met arg ser ala trp thr ala gln his leu
541/181	571/191
GAG TTG GCG GCG CCG AGG GCC GAG ATG GCA	GGA TGA CCG ATC GTC GGG GGC GGG AAC TCC
glu leu ala ala pro arg ala glu met ala	gly OPA arg ile val gly gly gly asn ser
601/201	631/211
CAG GCC GCC GGA CCG TCG CAA ACC CGT CGC	AAA CCC GTC GCA AAC CGT AAG GAG TCA TCC
gln ala ala gly pro ser gln thr arg arg	lys pro val ala asn arg lys glu ser ser
661/221	691/231
ATG AAG ACA GGC ACC GCG ACG ACG CGG CGC	AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG
met lys thr gly thr ala thr thr arg arg	arg leu leu ala val leu ile ala leu ala
721/241	751/251
TTG CCG GGG GCC GCC GTT GCG CTG CTG GCC	GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG
leu pro gly ala ala val ala leu leu ala	glu pro ser ala thr gly ala ser asp pro
781/261	811/271
TGC GCG GCC AGC GAA GTG GCG AGG ACG GTC	GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC
cys ala ala ser glu val ala arg thr val	gly ser val ala lys ser met gly asp tyr
841/281	871/291
CTG GAT TCA CAC CCA GAG ACC AAC CAG GTG	ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG
leu asp ser his pro glu thr asn gln val	met thr ala val leu gln gln gln val gly
901/301	931/311
CCG GGG TCG GTC GCA TCG CTG AAG GCC CAT	TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT C
pro gly ser val ala ser leu lys ala his	phe glu ala asn pro lys val ala ser asp

SEQ ID No.1A'

FIGURE 1A'

Insert of the clone containing DP428, other reading frame

2/1	32/11
ATC GCC TTT GAC GCC TAT TCG GTC GCG CAG CTT TTT GGC GAC GTC ACT GGT GCC CGC GTT	
ile ala phe asp ala tyr ser val ala gln leu phe gly asp val thr gly ala arg val	
62/21	92/31
GCG GGC GTC CAG CCG CAG CGA CAC CAC ATA CGG CCG GTC CGG TTG CGG GGG CCG TTG GGT	
ala gly val gln pro gln arg his his ile arg pro val arg leu arg gly pro leu gly	
122/41	152/51
GGG GTT GGG TGC CTC CGT CAC CCC AGG CAG TTC GCT GGC TAT TTG TCG CAG TAG CGC GAC	
gly val gly cys leu arg his pro arg gln phe ala gly tyr leu ser gln AMB arg asp	
182/61	212/71
GGC ATT GTC GAT GTC TTG GTA GCT AGC ATC CGG TCG GGG GGC CGC TAC CAG CGC CAG CGC	
gly ile val asp val leu val ala ser ile arg ser gly gly arg tyr gln arg gln arg	
242/81	272/91
CGG GGC TCC CCG GTC CGG GTA GTG CGC GTC GAG TTG GTC GTG GAC CAG CAA TGA CTG CGA	
arg gly ser pro val arg val val arg val glu leu val val asp gln gln OPA leu arg	
302/101	332/111
CCC GGC GAC TTC GAA ACC GCC ACC GGT TAG ATT CCC CGA CTG CGT CAT CGC CAG GTA AAC	
pro gly asp phe glu thr ala thr gly AMB ile pro arg leu arg his arg gln val asn	
362/121	392/131
CGC CGG CAC TAA CGC CAG CAA CCA ACC CGT GAA GAC CAA CCA ACG GCA CCT GCG CAG GTT	
arg arg his OCH arg gln gln pro thr arg glu asp gln pro thr ala pro ala gln val	
422/141	452/151
GCG GCT CAA CCG CAT CAT GAA CTG CTG GAT TTC GGA CTC CCC GTA CTC TCG CGC AGT GCG	
ala ala gln pro his his glu leu leu asp phe gly leu pro val leu ser arg ser ala	
482/161	512/171
TGC CCG CGA GCC TAC CGA AGA TCG CGT GCA TGC GTT CGG CGT GGA CCG CAC AGC ACC TGG	
cys pro arg ala tyr arg arg ser arg ala cys val arg arg gly pro his ser thr trp	
542/181	572/191
AGT TGG CGG CGC CGA GGG CCG AGA TGG CAG GAT GAC GGA TCG TCG GGG GCG GGA ACT CCC	
ser trp arg arg arg gly pro arg trp gln asp asp gly ser ser gly ala gly thr pro	
602/201	632/211
AGG CCG CCG GAC CGT CGC AAA CCC GTC GCA AAC CCG TCG CAA ACC GTA AGG AGT CAT CCA	
arg pro pro asp arg arg lys pro val ala asn pro ser gln thr val arg ser his pro	
662/221	692/231
TGA AGA CAG GCA CCG CGA CGA CGC GGC GCA GGC TGT TGG CAG TAC TGA TCG CCC TCG CGT	
OPA arg gln ala pro arg arg arg gly ala gly cys trp gln tyr OPA ser pro ser arg	
722/241	752/251
TGC CGG GGG CCG CCG TTG CGC TGC TGG CCG AAC CAT CAG CGA CCG GCG CGT CGG ACC CGT	
cys arg gly pro pro leu arg cys trp pro asn his gln arg pro ala arg arg thr arg	
782/261	812/271
GCG CGG CCA GCG AAG TGG CGA GGA CGG TCG GTT CGG TCG CCA AGT CGA TGG GCG ACT ACC	
ala arg pro ala lys trp arg gly arg ser val arg ser pro ser arg trp ala thr thr	
842/281	872/291
TGG ATT CAC ACC CAG AGA CCA ACC AGG TGA TGA CCG CGG TCT TGC AGC AGC AGG TAG GGC	
trp ile his thr gln arg pro thr arg OPA OPA pro arg ser cys ser ser arg AMB gly	
902/301	932/311
CGG GGT CGG TCG CAT CGC TGA AGG CCC ATT TCG AGG CGA ATC CCA AGG TCG CAT CGG ATC	
arg gly arg ser his arg OPA arg pro ile ser arg arg ile pro arg ser his arg ile	

SEQ ID No.1B'

FIGURE 1B'

Seq1C: Insert of the DP428 clone, other reading frame

3/1	33/11
TCG CCT TTG ACG CCT ATT CGG TCG CGC AGC	TTT TTG GCG ACG TCA CTG GTG CCC GCG TTG
ser pro leu thr pro ile arg ser arg ser	phe leu ala thr ser leu val pro ala leu
63/21	93/31
CGG GCG TCC AGC CGC AGC GAC ACC ACA TAC	GGC CGG TCC GGT TGC GGG GGC CGT TGG GTG
arg ala ser ser arg ser asp thr thr tyr	gly arg ser gly cys gly gly arg trp val
123/41	153/51
GGG TTG GGT GCC TCC GTC ACC CCA GGC AGT	TCG CTG GCT ATT TGT CGC AGT AGC GCG ACG
gly leu gly ala ser val thr pro gly ser	ser leu ala ile cys arg ser ser ala thr
183/61	213/71
GCA TTG TCG ATG TCT TGG TAG CTA GCA TCC	GGT CGG GGG GCC GCT ACC AGC GCC AGC GCC
ala leu ser met ser trp AMB leu ala ser	gly arg gly ala ala thr ser ala ser ala
243/81	273/91
GGG GCT CCC CGG TCC GGG TAG TGC GCG TCG	AGT TGG TCG TGG ACC AGC AAT GAC TGC GAC
gly ala pro arg ser gly AMB cys ala ser	ser trp ser trp thr ser asn asp cys asp
303/101	333/111
CCG GCG ACT TCG AAA CCG CCA CCG GTT AGA	TTC CCC GAC TGC GTC ATC GCC AGG TAA ACC
pro ala thr ser lys pro pro pro val arg	phe pro asp cys val ile ala arg OCH thr
363/121	393/131
GCC GGC ACT AAC GCC AGC AAC CAA CCC GTG	AAG ACC AAC CAA CGG CAC CTG CGC AGG TTG
ala gly thr asn ala ser asn gln pro val	lys thr asn gln arg his leu arg arg leu
423/141	453/151
CGG CTC AAC CGC ATC ATG AAC TGC TGG ATT	TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT
arg leu asn arg ile met asn cys trp ile	ser asp ser pro tyr ser arg ala val arg
483/161	513/171
GCC CGC GAG CCT ACC GAA GAT CGC GTG CAT	GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA
ala arg glu pro thr glu asp arg val his	ala phe gly val asp arg thr ala pro gly
543/181	573/191
GTT GGC GGC GCC GAG GGC CGA GAT GGC AGG	ATG ACG GAT CGT CGG GGG CGG GAA CTC CCA
val gly gly ala glu gly arg asp gly arg	met thr asp arg arg gly arg glu leu pro
603/201	633/211
GGC CGC CGG ACC GTC GCA AAC CCG TCG CAA	ACC CGT CGC AAA CCG TAA GGA GTC ATC CAT
gly arg arg thr val ala asn pro ser gln	thr arg arg lys pro OCH gly val ile his
663/221	693/231
GAA GAC AGG CAC CGC GAC GAC GCG GCG CAG	GCT GTT GGC AGT ACT GAT CGC CCT CGC GTT
glu asp arg his arg asp asp ala ala gln	ala val gly ser thr asp arg pro arg val
723/241	753/251
GCC GGG GGC CGC CGT TGC GCT GCT GGC CGA	ACC ATC AGC GAC CGG CGC GTC GGA CCC GTG
ala gly gly arg arg cys ala ala gly arg	thr ile ser asp arg arg val gly pro val
783/261	813/271
CGC GGC CAG CGA AGT GGC GAG GAC GGT CGG	TTC GGT CGC CAA GTC GAT GGG CGA CTA CCT
arg gly gln arg ser gly glu asp gly arg	phe gly arg gln val asp gly arg leu pro
843/281	873/291
GGA TTC ACA CCC AGA GAC CAA CCA GGT GAT	GAC CGC GGT CTT GCA GCA GCA GGT AGG GCC
gly phe thr pro arg asp gln pro gly asp	asp arg gly leu ala ala ala gly arg ala
903/301	933/311
GGG GTC GGT CGC ATC GCT GAA GGC CCA TTT	CGA GGC GAA TCC CAA GGT CGC ATC GGA TC
gly val gly arg ile ala glu gly pro phe	arg gly glu ser gln gly arg ile gly

SEQ ID No.1C'

FIGURE 1C'

Coding sequence DP428 identical to the Rv0203 predicted by Cole et al.
(Nature 393:537-544)

1/1	31/11
ATG AAG ACA GGC ACC GCG ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG	
Met lys thr gly thr ala thr thr arg arg	arg leu leu ala val leu ile ala leu ala
61/21	91/31
TTG CCG GGG GCC GCC GTT GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG	
leu pro gly ala ala val ala leu leu ala	glu pro ser ala thr gly ala ser asp pro
121/41	151/51
TGC GCG GCC AGC GAA GTG GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC	
cys ala ala ser glu val ala arg thr val	gly ser val ala lys ser met gly asp tyr
181/61	211/71
CTG GAT TCA CAC CCA GAG ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG	
leu asp ser his pro glu thr asn gln val	met thr ala val leu gln gln gln val gly
241/81	271/91
CCG GGG TCG GTC GCA TCG CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT	
pro gly ser val ala ser leu lys ala his	phe glu ala asn pro lys val ala ser asp
301/101	331/111
CTG CAC GCG CTT TCG CAA CCG CTG ACC GAT CTT TCG ACT CGG TGC TCG CTG CCG ATC AGC	
leu his ala leu ser gln pro leu thr asp	leu ser thr arg cys ser leu pro ile ser
361/121	391/131
GGC CTG CAG GCG ATC GGT TTG ATG CAG GCG GTG CAG GGC GCC CGC CGG TAG	
gly leu gln ala ile gly leu met gln ala val gln gly ala arg arg	AMB

SEQ ID No.1D

FIGURE 1D

ORF containing the DP428 sequence and forming part of seq1A'

1/1	31/11
TGA CGG ATC GTC GGG GGC GGG AAC TCC CAG GCC GCC GGA CCG TCG CAA ACC CGT CGC AAA	
OPA arg ile val gly gly gly asn ser gln	ala ala gly pro ser gln thr arg arg lys
61/21	91/31
CCC GTC GCA AAC CGT AAG GAG TCA TCC ATG AAG ACA GGC ACC GCG ACG ACG CGG CGC AGG	
pro val ala asn arg lys glu ser ser met	lys thr gly thr ala thr thr arg arg arg
121/41	151/51
CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG GGG GCC GCC GTT GCG CTG CTG GCC GAA	
leu leu ala val leu ile ala leu ala leu	pro gly ala ala val ala leu leu ala glu
181/61	211/71
CCA TCA GCG ACC GGC GCG TCG GAC CCG TGC GCG GCC AGC GAA GTG GCG AGG ACG GTC GGT	
pro ser ala thr gly ala ser asp pro cys	ala ala ser glu val ala arg thr val gly
241/81	271/91
TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT TCA CAC CCA GAG ACC AAC CAG GTG ATG	
ser val ala lys ser met gly asp tyr leu	asp ser his pro glu thr asn gln val met
301/101	331/111
ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG TCG GTC GCA TCG CTG AAG GCC CAT TTC	
thr ala val leu gln gln gln val gly pro	gly ser val ala ser leu lys ala his phe
361/121	391/131
GAG GCG AAT CCC AAG GTC GCA TCG GAT CTG CAC GCG CTT TCG CAA CCG CTG ACC GAT CTT	
glu ala asn pro lys val ala ser asp leu	his ala leu ser gln pro leu thr asp leu
421/141	451/151
TCG ACT CGG TGC TCG CTG CCG ATC AGC GGC CTG CAG GCG ATC GGT TTG ATG CAG GCG GTG	
ser thr arg cys ser leu pro ile ser gly leu gln ala ile gly leu met gln ala val	
481/161	
CAG GGC GCC CGC CGG TAG	
gln gly ala arg arg	AMB

SEQ ID No.1F

FIGURE 1F

491 CCGGTCGGGGGGCCGCTACCAGCGCCAGCGCCGGGGCTCCCCGGTCCGGGTA GTG CGC GTC GAG TTG GTC GTG 563
 1 V R V E L V V 7
 564 GAC CAG CAA TGA CTGCGACCCGGCGACTTCGAAACCGCCACCGGTTAGATTCCCCGACTGCGTCATCGCCAGGTAA 639
 8 D Q Q *
 640 ACCGCCGGCACTAACGCCAGCAACCAACCC GTG AAG ACC AAC CAA CGG CAC CTG CGC AGG TTG CGG 705
 1 V K T N Q R H L R R L R 12
 706 CTC AAC CGC ATC ATG AAC TGC TGG ATT TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT GCC 765
 13 L N R I M N C W I S D S P Y S R A V R A 32
 766 CGC GAG CCT ACC GAA GAT CGC GTG CAT GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA GTT 825
 33 R E P T E D R V H A F G V D R T A P G V 52
 826 GGC GGC GCC GAG GGC CGA GAT GGC AGG ATG ACG GAT CGT CGG GGG CGG GAA CTC CCA GGC 885
 53 G G A E G R D G R M T D R R G R E L P G 72
 886 CGC CGG ACC GTC GCA AAC CCG TCG CAA ACC CGT CGC AAA CCG TAA GGAGTCATCC ATG AAG 946
 73 R R T V A N P S Q T R R K P * xxxxxx M K 2
 947 ACA GGC ACC GCG ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG 1006
 3 T G T A T T R R R L L A V L I A L A L P 22
 1007 GGG GCC GCC GTT GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG TGC GCG 1066
 23 G A A V A L L A E P S A T G A S D P C A 42
 1067 GCC AGC GAA GTG GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT 1126
 43 A S E V A R T V G S V A K S M G D Y L D 62
 1127 TCA CAC CCA GAG ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG 1186
 63 S H P E T N Q V M T A V L Q Q Q V G P G 82
 1187 TCG GTC GCA TCG CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT CTG CAC 1246
 83 S V A S L K A H F E A N P K V A S D L H 102
 1247 GCG CTT TCG CAA CCG CTG ACC GAT CTT TCG ACT CGG TGC TCG CTG CCG ATC AGC GGC CTG 1306
 103 A L S Q P L T D L S T R C S L P I S G L 122
 1307 CAG GCG ATC GGT TTG ATG CAG GCG GTG CAG GGC GCC CGC CGG TAG ATG CCG GAC CGC CGC 1366
 123 Q A I G L M Q A V Q G A R R * M P D R R 5
 1367 CGG GTC CGG CGC AGT CGA CGT GAG GCA GCG GTC GCC TAC CGG GGC GGT GTC TCG CCG CCT 1426
 6 R V R R S R R E A A V A Y R G G V S P P 25
 1427 TCT GGT CGC AGG TCA GGG GTC GGC GCT GGA CCT TGC GGT GTG GTT TCG ACC GGG TCG TCG 1486
 26 S G R R S G V G A G P C G V V S T G S S 45
 1487 CAG GGT GTG CCC TGC GGT TGG ATG ACA AGT CGC AGG TTT GGA TCG GTT GGC GGG TCG CGA 1546
 46 Q G V P C G W M T S R R F G S V G G S R 65
 1547 TCG TTG T 1553
 66 S L 67

SEQ ID No.2

FIGURE 2

31/11
 TCG CCG GCT CGC GGA CGT AGA TAA TAG CTC ACC GTT GGA CGA CCT CGA CAG GGT CCT TTG
 ser pro ala arg gly arg arg OCH AMB leu thr val gly arg pro arg gln gly pro leu
 61/21 91/31
 TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GGT CAT CGC CTA AGG CTA CCG TTC
 OPA leu pro gly leu thr arg thr thr thr glu ser gly his arg leu arg leu pro phe
 121/41 151/51
 TGA CCT GGG GTG CGT GGG CGC CGA CGA GTG AGG CAG TCA TGT CTC AGG GCC CAC CGC CAC
 OPA pro gly val arg gly arg arg arg val arg gln ser cys leu arg ala his arg his
 181/61 211/71
 CTC GGT CGC CGG CAG TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC TTG TTC GTG TTG GTG
 leu gly arg arg gln cys gln his val gln met thr pro arg ser leu phe val leu val
 241/81 271/91
 TCG TGG TTG CGA CGA CTT GGC GCT GGT GAG CGC ACC CGC CGG CGT CGT GCC GCG CAT GCG
 ser trp leu arg arg leu gly ala gly glu arg thr arg arg arg arg ala ala his ala
 301/101
 GAT C
 asp

SEQ ID No.3A

FIGURE 3A

32/11
 CGC CGG CTC GCG GAC GTA GAT AAT AGC TCA CCG TTG GAC GAC CTC GAC AGG GTC CTT TGT
 arg arg leu ala asp val asp asn ser ser pro leu asp asp leu asp arg val leu cys
 62/21 92/31
 GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT CGG GTC ATC GCC TAA GGC TAC CGT TCT
 asp cys arg ala OPA arg gly arg pro gln ser arg val ile ala OCH gly tyr arg ser
 122/41 152/51
 GAC CTG GGG TGC GTG GGC GCC GAC GAG TGA GGC AGT CAT GTC TCA GGG CCC ACC GCC ACC
 asp leu gly cys val gly ala asp glu OPA gly ser his val ser gly pro thr ala thr
 182/61 212/71
 TCG GTC GCC GGC AGT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCT TGT TCG TGT TGG TGT
 ser val ala gly ser val ser met cys arg OPA leu his ala ala cys ser cys trp cys
 242/81 272/91
 CGT GGT TGC GAC GAC TTG GCG CTG GTG AGC GCA CCC GCC GGC GTC GTG CCG CGC ATG CGG
 arg gly cys asp asp leu ala leu val ser ala pro ala gly val val pro arg met arg
 302/101
 ATC
 ile

SEQ ID No.3B

FIGURE 3B

33/11
 GCC GGC TCG CGG ACG TAG ATA ATA GCT CAC CGT TGG ACG ACC TCG ACA GGG TCC TTT GTG
 ala gly ser arg thr AMB ile ile ala his arg trp thr thr ser thr gly ser phe val
 63/21
 ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGG TCA TCG CCT AAG GCT ACC GTT CTG
 thr ala gly leu asp ala asp asp his arg val gly ser ser pro lys ala thr val leu
 123/41
 ACC TGG GGT GCG TGG GCG CCG ACG AGT GAG GCA GTC ATG TCT CAG GGC CCA CCG CCA CCT
 thr trp gly ala trp ala pro thr ser glu ala val met ser gln gly pro pro pro pro
 183/61
 CGG TCG CCG GCA GTG TCA GCA TGT GCA GAT GAC TCC ACG CAG CTT GTT CGT GTT GGT GTC
 arg ser pro ala val ser ala cys ala asp asp ser thr gln leu val arg val gly val
 243/81
 213/71
 273/91
 GTG GTT GCG ACG ACT TGG CGC TGG TGA GCG CAC CCG CCG GCG TCG TGC CGC GCA TGC GGA
 val val ala thr thr trp arg trp OPA ala his pro pro ala ser cys arg ala cys gly

TC

SEQ ID No.3C

FIGURE 3C

31/11
 CCA ATT TTC CTT CGC GCC GTG CAA TAC CAT CTG CAA GAC CAG CGA CGG CCC GTG GTT GCG
 pro ile phe leu arg ala val gln tyr his leu gln asp gln arg arg pro val val ala
 61/21
 91/31
 GTC GCG CAG CTT GCG GAA ACC GGG TAT GGA CCC TGC CGT ACC GTT GTT GCC ACT TGA TGT
 val ala gln leu ala glu thr gly tyr gly pro cys arg thr val val ala thr OPA cys
 121/41
 151/51
 CGT CGC TCT CCA CCC GTC GGG GGG CGA AAG CCA TTC CGA CAC TGG GAT CCT CAA AAC GTC
 arg arg ser pro pro val gly gly arg lys pro phe arg his trp asp pro gln asn val
 181/61
 211/71
 GGC TGA GTG TCT GCA GGG CTC CGG GGA GCA GCC GAT CAT CAC CAT GTA CGA ACT GAA TAA
 gly OPA val ser ala gly leu arg gly ala ala asp his his his val arg thr glu OCH
 241/81
 271/91
 GTC CCC CGC GCG CGA CTT CCA GAC ATT TGT TGT GGT TTC GGT TGA GGC CGA GGC GAG GCT
 val pro arg ala arg leu pro asp ile cys cys gly phe gly OPA gly arg gly glu ala
 301/101
 331/111
 CAT TTC GCA GCA ACC GGT CTC CGG GTC GCA GCA TCG TTG CGG CGA TCG CGG CGC AGT CGT
 his phe ala ala thr gly leu arg val ala ala ser leu arg arg ser arg arg ser arg
 361/121
 CGG ACG AGT CGT CGT CAA CGA CCA CGA TC
 arg thr ser arg arg gln arg pro arg

SEQ ID No.4A

FIGURE 4A

32/11
 CAA TTT TCC TTC GCG CCG TGC AAT ACC ATC TGC AAG ACC AGC GAC GGC CCG TGG TTG CGG
 gln phe ser phe ala pro cys asn thr ile cys lys thr ser asp gly pro trp leu arg
 62/21
 TCG CGC AGC TTG CGG AAA CCG GGT ATG GAC CCT GCC GTA CCG TTG TTG CCA CTT GAT GTC
 ser arg ser leu arg lys pro gly met asp pro ala val pro leu leu pro leu asp val
 122/41
 GTC GCT CTC CAC CCG TCG GGG GGC GAA AGC CAT TCC GAC ACT GGG ATC CTC AAA ACG TCG
 val ala leu his pro ser gly gly glu ser his ser asp thr gly ile leu lys thr ser
 182/61
 GCT GAG TGT CTG CAG GGC TCC GGG GAG CAG CCG ATC ATC ACC ATG TAC GAA CTG AAT AAG
 ala glu cys leu gln gly ser gly glu gln pro ile ile thr met tyr glu leu asn lys
 242/81
 TCC CCC GCG CGC GAC TTC CAG ACA TTT GTT GTG GTT TCG GTT GAG GCC GAG GCG AGG CTC
 ser pro ala arg asp phe gln thr phe val val val ser val glu ala glu ala arg leu
 302/101
 ATT TCG CAG CAA CCG GTC TCC GGG TCG CAG CAT CGT TGC GGC GAT CGC GGC GCA GTC GTC
 ile ser gln gln pro val ser gly ser gln his arg cys gly asp arg gly ala val val
 362/121
 GGA CGA GTC GTC GTC AAC GAC CAC GAT C
 gly arg val val val asn asp his asp

SEQ ID No.4B

FIGURE 4B

33/11
 AAT TTT CCT TCG CGC CGT GCA ATA CCA TCT GCA AGA CCA GCG ACG GCC CGT GGT TGC GGT
 asn phe pro ser arg arg ala ile pro ser ala arg pro ala thr ala arg gly cys gly
 63/21
 CGC GCA GCT TGC GGA AAC CGG GTA TGG ACC CTG CCG TAC CGT TGT TGC CAC TTG ATG TCG
 arg ala ala cys gly asn arg val trp thr leu pro tyr arg cys cys his leu met ser
 123/41
 TCG CTC TCC ACC CGT CGG GGG GCG AAA GCC ATT CCG ACA CTG GGA TCC TCA AAA CGT CGG
 ser leu ser thr arg arg gly ala lys ala ile pro thr leu gly ser ser lys arg arg
 183/61
 CTG AGT GTC TGC AGG GCT CCG GGG AGC AGC CGA TCA TCA CCA TGT ACG AAC TGA ATA AGT
 leu ser val cys arg ala pro gly ser ser arg ser ser pro cys thr asn OPA ile ser
 243/81
 CCC CCG CGC GCG ACT TCC AGA CAT TTG TTG TGG TTT CGG TTG AGG CCG AGG CGA GGC TCA
 pro pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg arg gly ser
 303/101
 TTT CGC AGC AAC CGG TCT CCG GGT CGC AGC ATC GTT GCG GCG ATC GCG GCG CAG TCG TCG
 phe arg ser asn arg ser pro gly arg ser ile val ala ala ile ala ala gln ser ser
 363/121
 GAC GAG TCG TCG TCA ACG ACC ACG ATC
 asp glu ser ser ser thr thr thr ile

SEQ ID No.4C

FIGURE 4C

part of the nucleotide sequence of seq4A

```

1/1                               31/11
CCG CGC GCG ACT TCC AGA CAT TTG TTG TGG TTT CGG TTG AGG CCG AGG CGA GGC TCA TTT
pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg arg gly ser phe
61/21                               91/31
CGC AGC AAG CGG TCT CCG GGT CGC AGC ATC GTT GCG GCG ATC GCG GCG CAG TCG TCG GAC
arg ser lys arg ser pro gly arg ser ile val ala ala ile ala ala gln ser ser asp
121/41
GAG TCG TCG TCA ACG ACC ACG ATC
glu ser ser ser thr thr thr ile

```

SEQ ID No.4A'

FIGURE 4A'

```

1/1                               31/11
CGC GCG CGA CTT CCA GAC ATT TGT TGT GGT TTC GGT TGA GGC CGA GGC GAG GCT CAT TTC
arg ala arg leu pro asp ile cys cys gly phe gly OPA gly arg gly glu ala his phe
61/21                               91/31
GCA GCA AGC GGT CTC CGG GTC GCA GCA TCG TTG CGG CGA TCG CGG CGC AGT CGT CGG ACG
ala ala ser gly leu arg val ala ala ser leu arg arg ser arg arg ser arg thr
121/41
AGT CGT CGT CAA CGA CCA CGA TC
ser arg arg gln arg pro arg

```

SEQ ID No.4B'

FIGURE 4B'

```

1/1                               31/11
GCC GCG CGC GAC TTC CAG ACA TTT GTT GTG GTT TCG GTT GAG GCC GAG GCG AGG CTC ATT
ala ala arg asp phe gln thr phe val val val ser val glu ala glu ala arg leu ile
61/21                               91/31
TCG CAG CAA GCG GTC TCC GGG TCG CAG CAT CGT TGC GGC GAT CGC GGC GCA GTC GTC GGA
ser gln gln ala val ser gly ser gln his arg cys gly asp arg gly ala val val gly
121/41
CGA GTC GTC GTC AAC GAC CAC GAT C
arg val val val asn asp his asp

```

SEQ ID No.4C'

FIGURE 4C'

11/185

ORF according to Cole et al. (Nature 393:537-544) and containing the sequence Seq 4A'

1/1 31/11
tga ata agt ccg ccg cgc gcg act tcc aga cat ttg ttg tgg ttt cgg ttg agg ccg agg
OPA ile ser pro pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg
61/21 91/31
cga ggc tca ttt cgc agc aag cgg tct ccg ggt cgc agc atc gtt gcg gcg atc gcg gcg
arg gly ser phe arg ser lys arg ser pro gly arg ser ile val ala ala ile ala ala
121/41 151/51
cag tcg tcg gac gag tcg tcg tca acg acc acg atc tcg aac tcg acg ccc tcc tgt tcg
gln ser ser asp glu ser ser ser thr thr thr ile ser asn ser thr pro ser cys ser
181/61 211/71
agg atg cta cgc aga cag cgc tcg atg gtg gcg ccg ttg ttg tac atc ggg atg cac acc
arg met leu arg arg gln arg ser met val ala pro leu leu tyr ile gly met his thr
241/81 271/91
gag ata agc ggt ttc gcc ggg ttc acc gat acc acg ctt gat gca tca cca ggc acc aca
glu ile ser gly phe ala gly phe thr asp thr thr leu asp ala ser pro gly thr thr
301/101
tgg cga ctc aga gac tag
trp arg leu arg asp AMB

SEQ ID No.4F

FIGURE 4F

sequence upstream of seq4A' and fused with seq4A'

1/1 31/11
GCA ACC TAC CAG CAG AGC CAG GGG CTC ACA GGA CCT AAA GGA GTA GCG CCC ATG GCT GAT
ala thr tyr gln gln ser gln gly leu thr gly pro lys gly val ala pro met ala asp

C

SEQ ID No.4J

FIGURE 4J

seq4J' in another reading frame

1/1 31/11
ACG CAA CCT ACC AGC AGA GCC AGG GGC TCA CAG GAC CTA AAG GAG TAG CGC CCA TGG CTG
thr gln pro thr ser arg ala arg gly ser gln asp leu lys glu AMB arg pro trp leu
61/21
ATC
ile

SEQ ID N°4K

FIGURE 4K

seq 4J' in the third reading frame

1/1 31/11
CGC AAC CTA CCA GCA GAG CCA GGG GCT CAC AGG ACC TAA AGG AGT AGC GCC CAT GGC TGA
arg asn leu pro ala glu pro gly ala his arg thr OCH arg ser ser ala his gly OPA

TC

SEQ ID No.4L

FIGURE 4L

sequence Rv2050 predicted by Cole et al. (Nature 393:537-544) and containing seq4J

```

1/1                                     31/11
ATG GCT GAT CGT GTC CTG AGG GGC AGT CGC CTC GGA GCC GTG AGC TAT GAG ACC GAC CGC
Met ala asp arg val leu arg gly ser arg leu gly ala val ser tyr glu thr asp arg
61/21                                     91/31
AAC CAC GAC CTG GCG CCG CGC CAG ATC GCG CGG TAC CGC ACC GAC AAC GGC GAG GAG TTC
asn his asp leu ala pro arg gln ile ala arg tyr arg thr asp asn gly glu glu phe
121/41                                     151/51
GAA GTC CCG TTC GCC GAT GAC GCC GAG ATC CCC GGC ACC TGG TTG TGC CGC AAC GGC ATG
glu val pro phe ala asp asp ala glu ile pro gly thr trp leu cys arg asn gly met
181/61                                     211/71
GAA GGC ACC CTG ATC GAG GGC GAC CTG CCC GAG CCG AAG AAG GTT AAG CCG CCC CGG ACG
glu gly thr leu ile glu gly asp leu pro glu pro lys lys val lys pro pro arg thr
241/81                                     271/91
CAC TGG GAC ATG CTG CTG GAG CGC CGT TCC ATC GAA GAA CTC GAA GAG TTA CTT AAG GAG
his trp asp met leu leu glu arg arg ser ile glu glu leu glu glu leu leu lys glu
301/101                                     331/111
CGC CTC GAG CTC ATT CGG TCA CGT CGG CGC GGC TGA
arg leu glu leu ile arg ser arg arg arg gly OPA

```

SEQ ID No.4M

FIGURE 4M

ORF according to Cole et al. (Nature 393:537-544) and containing the sequence Rv2050

```

1/1                                     31/11
TAG TCC GCC CGG GTG TCC GAT CCC GGT ATC ATT GAT GGT CGC GCC GCG CGC GTC GCG TGC
AMB ser ala arg val ser asp pro gly ile ile asp gly arg ala ala arg val ala cys
61/21                                     91/31
CGG GAA CTA CGC AGA CGG CCG CAG CGT TTG CCA ACC GGA GCC AGT CGC CAG TAC GCA ACC
arg glu leu arg arg arg pro gln arg leu pro thr gly ala ser arg gln tyr ala thr
121/41                                     151/51
TAC CAG CAG AGC CCA GGG CTC ACA GGA CCT AAA GGA GTA GCG CCC ATG GCT GAT CGT GTC
tyr gln gln ser pro gly leu thr gly pro lys gly val ala pro met ala asp arg val
181/61                                     211/71
CTG AGG GGC AGT CGC CTC GGA GCC GTG AGC TAT GAG ACC GAC CGC AAC CAC GAC CTG GCG
leu arg gly ser arg leu gly ala val ser tyr glu thr asp arg asn his asp leu ala
241/81                                     271/91
CCG CGC CAG ATC GCG CGG TAC CGC ACC GAC AAC GGC GAG GAG TTC GAA GTC CCG TTC GCC
pro arg gln ile ala arg tyr arg thr asp asn gly glu glu phe glu val pro phe ala
301/101                                     331/111
GAT GAC GCC GAG ATC CCC GGC ACC TGG TTG TGC CGC AAC GGC ATG GAA GGC ACC CTG ATC
asp asp ala glu ile pro gly thr trp leu cys arg asn gly met glu gly thr leu ile
361/121                                     391/131
GAG GGC GAC CTG CCC GAG CCG AAG AAG GTT AAG CCG CCC CGG ACG CAC TGG GAC ATG CTG
glu gly asp leu pro glu pro lys lys val lys pro pro arg thr his trp asp met leu
421/141                                     451/151
CTG GAG CGC CGT TCC ATC GAA GAA CTC GAA GAG TTA CTT AAG GAG CGC CTC GAG CTC ATT
leu glu arg arg ser ile glu glu leu glu glu leu leu lys glu arg leu glu leu ile
481/161
CGG TCA CGT CGG CGC GGC TGA
arg ser arg arg arg gly OPA

```

SEQ ID No.4N

FIGURE 4

13/185

31/11
GAT CGC GGT CAA CGA GGC CGA ATA CGG CGA GAT GTG GGC CCA AGA CGC CGC CGC GAT GTT
asp arg gly gln arg gly arg ile arg arg asp val gly pro arg arg arg arg asp val
61/21 91/31
TGG CTA CGC CGC GGC GAC GGC GAC GGC GAC GGC GAC GTT GCT GCC GTT CGA GGA GGC GCC
trp leu arg arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala
121/41 151/51
GGA GAT GAC CAG CGC GGG TGG GCT CCT CGA GCA GGC CGC CGC GGT CGA GGA GGC CTC CGA
gly asp asp gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg
181/61 211/71
CAC CGC CGC GGC GAA CCA GTT GAT GAA CAA TGT GCC CCA GGC GCT GCA ACA GCT GGC CCA
his arg arg gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro
241/81 271/91
GCC CAC GCA GGC CAC CAC GCC TTC TTC CAA GCT GGG TGG CCT GTG GAA GAC GGT CTC GCC
ala his ala gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala
301/101
GCA TCG GTC GCC GAT C
ala ser val ala asp

SEQ ID No.5A

FIGURE 5A

32/11
ATC GCG GTC AAC GAG GCC GAA TAC GGC GAG ATG TGG GCC CAA GAC GCC GCC GCG ATG TTT
ile ala val asn glu ala glu tyr gly glu met trp ala gln asp ala ala ala met phe
62/21 92/31
GGC TAC GCC GCG GCG ACG GCG ACG GCG ACG GCG ACG TTG CTG CCG TTC GAG GAG GCG CCG
gly tyr ala ala ala thr ala thr ala thr ala thr leu leu pro phe glu glu ala pro
122/41 152/51
GAG ATG ACC AGC GCG GGT GGG CTC CTC GAG CAG GCC GCC GCG GTC GAG GAG GCC TCC GAC
glu met thr ser ala gly gly leu leu glu gln ala ala ala val glu glu ala ser asp
182/61 212/71
ACC GCC GCG GCG AAC CAG TTG ATG AAC AAT GTG CCC CAG GCG CTG CAA CAG CTG GCC CAG
thr ala ala ala asn gln leu met asn asn val pro gln ala leu gln gln leu ala gln
242/81 272/91
CCC ACG CAG GGC ACC ACG CCT TCT TCC AAG CTG GGT GGC CTG TGG AAG ACG GTC TCG CCG
pro thr gln gly thr thr pro ser ser lys leu gly gly leu trp lys thr val ser pro
302/101
CAT CGG TCG CCG ATC
his arg ser pro ile

SEQ ID No.5B

FIGURE 5B

33/11
 TCG CGG TCA ACG AGG CCG AAT ACG GCG AGA TGT GGG CCC AAG ACG CCG CCG CGA TGT TTG
 ser arg ser thr arg pro asn thr ala arg cys gly pro lys thr pro pro arg cys leu
 63/21
 GCT ACG CCG CGG CGA CGG CGA CGG CGA CGG CGA CGT TGC TGC CGT TCG AGG AGG CGC CGG
 ala thr pro arg arg arg arg arg arg arg arg arg arg cys cys arg ser arg arg arg arg
 123/41
 AGA TGA CCA GCG CGG GTG GGC TCC TCG AGC AGG CCG CCG CGG TCG AGG AGG CCT CCG ACA
 arg OPA pro ala arg val gly ser ser ser arg pro pro arg ser arg arg pro pro thr
 183/61
 CCG CCG CGG CGA ACC AGT TGA TGA ACA ATG TGC CCC AGG CGC TGC AAC AGC TGG CCC AGC
 pro pro arg arg thr ser OPA OPA thr met cys pro arg arg cys asn ser trp pro ser
 243/81
 CCA CGC AGG GCA CCA CGC CTT CTT CCA AGC TGG GTG GCC TGT GGA AGA CGG TCT CGC CGC
 pro arg arg ala pro arg leu leu pro ser trp val ala cys gly arg arg ser arg arg
 303/101
 ATC GGT CGC CGA TC
 ile gly arg arg

SEQ ID No.5C

FIGURE 5C

part of the nucleotide sequence Seq 5A

1/1
 CGC CGC GGC GAC GGC GAC GGC GAC GGC GAC GTT GCT GCC GTT CGA GGA GGC GCC GGA GAT
 arg arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala gly asp
 61/21
 GAC CAG CGC GGG TGG GCT CCT CGA GCA GGC CGC CGC GGT CGA GGA GGC CTC CGA CAC CGC
 asp gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg his arg
 121/41
 CGC GGC GAA CCA GTT GAT GAA CAA TGT GCC CCA GGC GCT GCA ACA GCT GGC CCA GCC CAC
 arg gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro ala his
 181/61
 GCA GGG CAC CAC GCC TTC TTC CAA GCT GGG TGG CCT GTG GAA GAC GGT CTC GCC GCA TCG
 ala gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala ala ser
 241/81
 GTC GCC GAT C
 val ala asp

SEQ ID No.5A'

FIGURE 5A'

1/1 31/11
TAC GCC GCG GCG ACG GCG ACG GCG ACG GCG ACG TTG CTG CCG TTC GAG GAG GCG CCG GAG
tyr ala ala ala thr ala thr ala thr ala thr leu leu pro phe glu glu ala pro glu
61/21 91/31
ATG ACC AGC GCG GGT GGG CTC CTC GAG CAG GCC GCC GCG GTC GAG GAG GCC TCC GAC ACC
met thr ser ala gly gly leu leu glu gln ala ala ala val glu glu ala ser asp thr
121/41 151/51
GCC GCG GCG AAC CAG TTG ATG AAC AAT GTG CCC CAG GCG CTG CAA CAG CTG GCC CAG CCC
ala ala ala asn gln leu met asn asn val pro gln ala leu gln gln leu ala gln pro
181/61 211/71
ACG CAG GGC ACC ACG CCT TCT TCC AAG CTG GGT GGC CTG TGG AAG ACG GTC TCG CCG CAT
thr gln gly thr thr pro ser ser lys leu gly gly leu trp lys thr val ser pro his
241/81
CGG TCG CCG ATC
arg ser pro ile

SEQ ID No.5B'

FIGURE 5B'

1/1 31/11
ACG CCG CGG CGA CGG CGA CGG CGA CGG CGA CGT TGC TGC CGT TCG AGG AGG CGC CGG AGA
thr pro arg arg arg arg arg arg arg arg arg arg arg cys cys arg ser arg arg arg arg
61/21 91/31
TGA CCA GCG CGG GTG GGC TCC TCG AGC AGG CCG CCG CGG TCG AGG AGG CCT CCG ACA CCG
OPA pro ala arg val gly ser ser ser arg pro pro arg ser arg arg pro pro thr pro
121/41 151/51
CCG CGG CGA ACC AGT TGA TGA ACA ATG TGC CCC AGG CGC TGC AAC AGC TGG CCC AGC CCA
pro arg arg thr ser OPA OPA thr met cys pro arg arg cys asn ser trp pro ser pro
181/61 211/71
CGC AGG GCA CCA CGC CTT CTT CCA AGC TGG GTG GCC TGT GGA AGA CGG TCT CGC CGC ATC
arg arg ala pro arg leu leu pro ser trp val ala cys gly arg arg ser arg arg ile
241/81
GGT CGC CGA TC
gly arg arg

SEQ ID No.5C'

FIGURE 5C'

ORF predicted by Cole et al. (Nature 393:537-544) and containing seq5A'

```

1/1                                31/11
tga act gat gat tct gat agc gac caa cct ctt ggg gca aaa cac ccc ggc gat cgc ggt
OPA thr asp asp ser asp ser asp gln pro leu gly ala lys his pro gly asp arg gly
61/21                                91/31
caa cga ggc cga ata cgg cga gat gtg ggc cca aga cgc cgc cgc gat gtt tgg cta cgc
gln arg gly arg ile arg arg asp val gly pro arg arg arg arg asp val trp leu arg
121/41                                151/51
cgc ggc gac ggc gac ggc gac ggc gac gtt gct gcc gtt cga gga ggc gcc gga gat gac
arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala gly asp asp
181/61                                211/71
cag cgc ggg tgg gct cct cga gca ggc cgc cgc ggt cga gga ggc ctc cga cac cgc cgc
gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg his arg arg
241/81                                271/91
ggc gaa cca gtt gat gaa caa tgt gcc cca ggc gct gca aca gct ggc cca gcc cac gca
gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro ala his ala
301/101                                331/111
ggg cac cac gcc ttc ttc caa gct ggg tgg cct gtg gaa gac ggt ctc gcc gca tcg gtc
gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala ala ser val
361/121                                391/131
gcc gat cag caa cat ggt gtc gat ggc caa caa cca cat gtc gat gac caa ctc ggg tgt
ala asp gln gln his gly val asp gly gln gln pro his val asp asp gln leu gly cys
421/141                                451/151
gtc gat gac caa cac ctt gag ctc gat gtt gaa ggg ctt tgc tcc ggc ggc ggc cgc cca
val asp asp gln his leu glu leu asp val glu gly leu cys ser gly gly gly arg pro
481/161                                511/171
ggc cgt gca aac cgc ggc gca aaa cgg ggt ccg ggc gat gag ctc gct ggg cag ctc gct
gly arg ala asn arg gly ala lys arg gly pro gly asp glu leu ala gly gln leu ala
541/181                                571/191
ggg ttc ttc ggg tct ggg cgg tgg ggt ggc cgc caa ctt ggg tgc ggc ggc ctc ggt cgg
gly phe phe gly ser gly arg trp gly gly arg gln leu gly ser gly gly leu gly arg
601/201                                631/211
ttc gtt gtc ggt gcc gca ggc ctg ggc cgc ggc caa cca ggc agt cac ccc ggc ggc gcg
phe val val gly ala ala gly leu gly arg gly gln pro gly ser his pro gly gly ala
661/221                                691/231
ggc gct gcc gct gac cag cct gac cag cgc cgc gga aag agg gcc cgg gca gat gct ggg
gly ala ala ala asp gln pro asp gln arg arg gly lys arg ala arg ala asp ala gly
721/241                                751/251
cgg gct gcc ggt ggg gca gat ggg cgc cag ggc cgg tgg tgg gct cag tgg tgt gct gcg
arg ala ala gly gly ala asp gly arg gln gly arg trp trp ala gln trp cys ala ala
781/261                                811/271
tgt tcc gcc gcg acc cta tgt gat gcc gca ttc tcc ggc ggc cgg cta gga gag ggg gcg
cys ser ala ala thr leu cys asp ala ala phe ser gly gly arg leu gly glu gly ala
841/281
cag act gtc gtt att tga
gln thr val val ile OPA

```

SEQ ID No.5F

FIGURE 5F

sequence Rv1196 predicted by Cole et al. (Nature 393:537-544) and capable of encoding an ORF fused with Seq5A'

```

1/1                                     31/11
atg gtg gat ttc ggg gcg tta cca ccg gag atc aac tcc gcg agg atg tac gcc ggc ccg
Met val asp phe gly ala leu pro pro glu ile asn ser ala arg met tyr ala gly pro
61/21                                 91/31
ggg tcc gcc tcg ctg gtg gcc gcg gct cag atg tgg gac agc gtg gcg agt gac ctg ttt
gly ser ala ser leu val ala ala ala gln met trp asp ser val ala ser asp leu phe
121/41                               151/51
tcg gcc gcg tcg gcg ttt cag tcg gtg gtc tgg ggt ctg acg gtg ggg tcg tgg ata ggt
ser ala ala ser ala phe gln ser val val trp gly leu thr val gly ser trp ile gly
181/61                               211/71
tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tcg ccg tat gtg gcg tgg atg agc gtc acc
ser ser ala gly leu met val ala
ala ala ser pro tyr val ala trp met ser val thr
241/81                               271/91
gcg ggg cag gcc gag ctg acc gcc gcc cag gtc ccg gtt gct gcg gcg gcc tac gag acg
ala gly gln ala glu leu thr ala ala gln val arg val ala ala ala tyr glu thr
301/101                             331/111
gcg tat ggg ctg acg gtg ccc ccg ccg gtg atc gcc gag aac cgt gct gaa ctg atg att
ala tyr gly leu thr val pro pro pro val ile ala glu asn arg ala glu leu met ile
361/121                             391/131
ctg ata gcg acc aac ctc ttg ggg caa aac acc ccg gcg atc gcg gtc aac gag gcc gaa
leu ile ala thr asn leu leu gly gln asn thr pro ala ile ala val asn glu ala glu
421/141                             451/151
tac gcc gag atg tgg gcc caa gac gcc gcc gcg atg ttt ggc tac gcc gcg gcg acg gcg
tyr gly glu met trp ala gln asp ala ala ala met phe gly tyr ala ala ala thr ala
481/161                             511/171
acg gcg acg gcg acg ttg ctg ccg ttc gag gag gcg ccg gag atg acc agc gcg ggt ggg
thr ala thr ala thr leu leu pro phe glu glu ala pro glu met thr ser ala gly gly
541/181                             571/191
ctc ctc gag cag gcc gcc gcg gtc gag gag gcc tcc gac acc gcc gcg gcg aac cag ttg
leu leu glu gln ala ala ala val glu glu ala ser asp thr ala ala ala asn gln leu
601/201                             631/211
atg aac aat gtg ccc cag gcg ctg caa cag ctg gcc cag ccc acg cag ggc acc acg cct
met asn asn val pro gln ala leu gln gln leu ala gln pro thr gln gly thr thr pro
661/221                             691/231
tct tcc aag ctg ggt ggc ctg tgg aag acg gtc tcg ccg cat cgg tcg ccg atc agc aac
ser ser lys leu gly gly leu trp lys thr val ser pro his arg ser pro ile ser asn
721/241                             751/251
atg gtg tcg atg gcc aac aac cac atg tcg atg acc aac tcg ggt gtg tcg atg acc aac
met val ser met ala asn asn his met ser met thr asn ser gly val ser met thr asn
781/261                             811/271
acc ttg agc tcg atg ttg aag ggc ttt gct ccg gcg gcg gcc gcc cag gcc gtg caa acc
thr leu ser ser met leu lys gly phe ala pro ala ala ala ala gln ala val gln thr
841/281                             871/291
gcg gcg caa aac ggg gtc ccg gcg atg agc tcg ctg ggc agc tcg ctg ggt tct tcg ggt
ala ala gln asn gly val arg ala met ser ser leu gly ser ser leu gly ser ser gly
901/301                             931/311
ctg ggc ggt ggg gtg gcc gcc aac ttg ggt ccg gcg gcc tcg gtc ggt tcg ttg tcg gtg
leu gly gly gly val ala ala asn leu gly arg ala ala ser val gly ser leu ser val
961/321                             991/331
ccg cag gcc tgg gcc gcg gcc aac cag gca gtc acc ccg gcg gcg ccg gcg ctg ccg ctg
pro gln ala trp ala ala ala asn gln ala val thr pro ala ala arg ala leu pro leu
1021/341                           1051/351
acc agc ctg acc agc gcc gcg gaa aga ggg ccc ggg cag atg ctg ggc ggg ctg ccg gtg
thr ser leu thr ser ala ala glu arg gly pro gly gln met leu gly gly leu pro val
1081/361                           1111/371
ggg cag atg ggc gcc agg gcc ggt ggt ggg ctc agt ggt gtg ctg cgt gtt ccg ccg cga
gly gln met gly ala arg ala gly gly gly leu ser gly val leu arg val pro pro arg
1141/381                           1171/391
ccc tat gtg atg ccg cat tct ccg gcg gcc ggc tag
pro tyr val met pro his ser pro ala ala gly AMB

```

SEQ ID No.5R

FIGURE 5R

REPLACEMENT SHEET (RULE 26)

Seq 5P: ORF according to Cole et al. (Nature 393:537-544) and containing the sequence Rv1196

```

1/1                                     31/11
tag gga cac gta atg gtg gat ttc ggg gcg tta cca ccg gag atc aac tcc gcg agg atg
AMB gly his val met val asp phe gly ala leu pro pro glu ile asn ser ala arg met
61/21                                     91/31
tac gcc ggc ccg ggt tgc gcc tgc ctg gtg gcc gcg gct cag atg tgg gac agc gtg gcg
tyr ala gly pro gly ser ala ser leu val ala ala ala gln met trp asp ser val ala
121/41                                     151/51
agt gac ctg ttt tgc gcc gcg tgc gcg ttt cag tgc gtg gtc tgg ggt ctg acg gtg ggg
ser asp leu phe ser ala ala ser ala phe gln ser val val trp gly leu thr val gly
181/61                                     211/71
tcg tgg ata ggt tgc tgc gcg ggt ctg atg gtg gcg gcg gcc tgc ccg tat gtg gcg tgg
ser trp ile gly ser ser ala gly leu met val ala ala ala ser pro tyr val ala trp
241/81                                     271/91
atg agc gtc acc gcg ggg cag gcc gag ctg acc gcc gcc cag gtc ccg gtt gct gcg gcg
met ser val thr ala gly gln ala glu leu thr ala ala gln val arg val ala ala ala
301/101                                    331/111
gcc tac gag acg gcg tat ggg ctg acg gtg ccc ccg ccg gtg atc gcc gag aac cgt gct
ala tyr glu thr ala tyr gly leu thr val pro pro pro val ile ala glu asn arg ala
361/121                                    391/131
gaa ctg atg att ctg ata gcg acc aac ctc ttg ggg caa aac acc ccg gcg atc gcg gtc
glu leu met ile leu ile ala thr asn leu leu gly gln asn thr pro ala ile ala val
421/141                                    451/151
aac gag gcc gaa tac ggc gag atg tgg gcc caa gac gcc gcc gcg atg ttt ggc tac gcc
asn glu ala glu tyr gly glu met trp ala gln asp ala ala ala met phe gly tyr ala
481/161                                    511/171
gcg gcg acg gcg acg gcg acg gcg acg ttg ctg ccg ttc gag gag gcg ccg gag atg acc
ala ala thr ala thr ala thr ala thr leu leu pro phe glu glu ala pro glu met thr
541/181                                    571/191
agc gcg ggt ggg ctc ctc gag cag gcc gcc ccg gtc gag gag gcc tcc gac acc gcc gcg
ser ala gly gly leu leu glu gln ala ala ala val glu glu ala ser asp thr ala ala
601/201                                    631/211
gcg aac cag ttg atg aac aat gtg ccc cag gcg ctg caa cag ctg gcc cag ccc acg cag
ala asn gln leu met asn asn val pro gln ala leu gln gln leu ala gln pro thr gln
661/221                                    691/231
ggc acc acg cct tct tcc aag ctg ggt ggc ctg tgg aag acg gtc tgc ccg cat cgg tgc
gly thr thr pro ser ser lys leu gly gly leu trp lys thr val ser pro his arg ser
721/241                                    751/251
ccg atc agc aac atg gtg tgc atg gcc aac aac cac atg tgc atg acc aac tgc ggt gtg
pro ile ser asn met val ser met ala asn asn his met ser met thr asn ser gly val
781/261                                    811/271
tcg atg acc aac acc ttg agc tgc atg ttg aag ggc ttt gct ccg gcg gcg gcc gcc cag
ser met thr asn thr leu ser ser met leu lys gly phe ala pro ala ala ala ala gln
841/281                                    871/291
gcc gtg caa acc gcg gcg caa aac ggg gtc ccg gcg atg agc tgc ctg gcc agc tgc ctg
ala val gln thr ala ala gln asn gly val arg ala met ser ser leu gly ser ser leu
901/301                                    931/311
ggt tct tgc ggt ctg ggc ggt ggg gtg gcc gcc aac ttg ggt ccg gcg gcc tgc gtc ggt
gly ser ser gly leu gly gly gly val ala ala asn leu gly arg ala ala ser val gly
961/321                                    991/331
tcg ttg tgc gtg ccg cag gcc tgg gcc gcg gcc aac cag gca gtc acc ccg gcg gcg ccg
ser leu ser val pro gln ala trp ala ala ala asn gln ala val thr pro ala ala arg
1021/341                                    1051/351
gcg ctg ccg ctg acc agc ctg acc agc gcc gcg gaa aga ggg ccc ggg cag atg ctg gcc
ala leu pro leu thr ser leu thr ser ala ala glu arg gly pro gly gln met leu gly
1081/361 1111/371
ggg ctg ccg gtg ggg cag atg ggc gcc agg gcc ggt ggt ggg ctc agt ggt gtg ctg cgt
gly leu pro val gly gln met gly ala arg ala gly gly gly leu ser gly val leu arg
1141/381                                    1171/391
ggt ccg ccg cga ccc tat gtg atg ccg cat tct ccg gcg gcc gcc tag
val pro pro arg pro tyr val met pro his ser pro ala ala gly AMB

```

SEQ ID No.5P

FIGURE 5P

REPLACEMENT SHEET (RULE 26)

31/11
 GGA TCC TGA TGC AAG TGG TCC GGG ATT TGT CGG CAG CCA CCG CGG TCC CGT CGA CCA ACG
 gly ser OPA cys lys trp ser gly ile cys arg gln pro arg arg ser arg arg pro thr
 61/21
 TTG GTG CAT CCG GGC TGC GAG CAT GCA CGC ACC GAC CAG CGC GGC GAG CGC GGC TAG CTG
 leu val his pro gly cys glu his ala arg thr asp gln arg gly glu arg gly AMB leu
 121/41
 CTT GCC CAC TGT TCC TCC CTG CCG GCA CCA TGT GCG ACA AGC TTA AGC GCA GCA GTA CCG
 leu ala his cys ser ser leu pro ala pro cys ala thr ser leu ser ala ala val pro
 181/61
 GCG GTG CCT GGG CAT CCA GCA AAA CGG GGA GCT CAA GAA CGA TTC ATG AAC GAG GGG TCG
 ala val pro gly his pro ala lys arg gly ala gln glu arg phe met asn glu gly sec
 241/81
 TCA CCA ACG TCG AAA CCG ACG GTT GCC AGC CGG CCC ACG ATA TTG CGT GCT CGA GGG TCC
 ser pro thr ser lys pro thr val ala ser arg pro thr ile leu arg ala arg gly ser
 301/101
 GCT GTA CCC TCA CCG AAC GTG AGT CCC ACA CCG CGG AGG CGG GCG ACT CTG GCG TCG TTA
 ala val pro ser pro asn val ser pro thr pro arg arg arg ala thr leu ala ser leu
 361/121
 GCA GCC GAG CTC AAG GTG TCC CGC ACC ACT GTC TCG AAT GCT TTT AAC CGA CCG GAT CCA
 ala ala glu leu lys val ser arg thr thr val ser asn ala phe asn arg pro asp pro
 421/141
 GAA GGA GAA GAT C
 glu gly glu asp

SEQ ID No.6A

FIGURE 6A

32/11
 GAT CCT GAT GCA AGT GGT CCG GGA TTT GTC GGC AGC CAC GGC GGT CCC GTC GAC CAA CGT
 asp pro asp ala ser gly pro gly phe val gly ser his gly gly pro val asp gln arg
 62/21
 TGG TGC ATC CGG GCT GCG AGC ATG CAC GCA CCG ACC AGC GCG GCG AGC GCG GCT AGC TGC
 trp cys ile arg ala ala ser met his ala pro thr ser ala ala ser ala ala ser cys
 122/41
 TTG CCC ACT GTT CCT CCC TGC CGG CAC CAT GTG CGA CAA GCT TAA GCG CAG CAG TAC CGG
 leu pro thr val pro pro cys arg his his val arg gln ala OCH ala gln gln tyr arg
 182/61
 CGG TGC CTG GGC ATC CAG CAA AAC GGG GAG CTC AAG AAC GAT TCA TGA ACG AGG GGT CGT
 arg cys leu gly ile gln gln asn gly glu leu lys asn asp ser OPA thr arg gly arg
 242/81
 CAC CAA CGT CGA AAC CGA CGG TTG CCA GCC GGC CCA CGA TAT TGC GTG CTC GAG GGT CCG
 his gln arg arg asn arg arg leu pro ala gly pro arg tyr cys val leu glu gly pro
 302/101
 CTG TAC CCT CAC CGA ACG TGA GTC CCA CAC CGC GGA GGC GGG CGA CTC TGG CGT CGT TAG
 leu tyr pro his arg thr OPA val pro his arg gly gly gly arg leu trp arg arg AMB
 362/121
 CAG CCG AGC TCA AGG TGT CCC GCA CCA CTG TCT CGA ATG CTT TTA ACC GAC CGG ATC CAG
 gln pro ser ser arg cys pro ala pro leu ser arg met leu leu thr asp arg ile gln
 422/141
 AAG GAG AAG ATC
 lys glu lys ile

SEQ ID No.6B

FIGURE 6B

33/11
 ATC CTG ATG CAA GTG GTC CGG GAT TTG TCG GCA GCC ACG GCG GTC CCG TCG ACC AAC GTT
 ile leu met gln val val arg asp leu ser ala ala thr ala val pro ser thr asn val
 63/21
 GGT GCA TCC GGG CTG CGA GCA TGC ACG CAC CGA CCA GCG CGG CGA GCG CGG CTA GCT GCT
 gly ala ser gly leu arg ala cys thr his arg pro ala arg arg ala arg leu ala ala
 123/41
 TGC CCA CTG TTC CTC CCT GCC GGC ACC ATG TGC GAC AAG CTT AAG CGC AGC AGT ACC GGC
 cys pro leu phe leu pro ala gly thr met cys asp lys leu lys arg scr ser thr gly
 183/61
 GGT GCC TGG GCA TCC AGC AAA ACG GGG AGC TCA AGA ACG ATT CAT GAA CGA GGG GTC GTC
 gly ala trp ala ser ser lys thr gly ser ser arg thr ile his glu arg gly val val
 243/81
 ACC AAC GTC GAA ACC GAC GGT TGC CAG CCG GCC CAC GAT ATT GCG TGC TCG AGG GTC CGC
 thr asn val glu thr asp gly cys gln pro ala his asp ile ala cys ser arg val arg
 303/101
 TGT ACC CTC ACC GAA CGT GAG TCC CAC ACC GCG GAG GCG GGC GAC TCT GGC GTC GTT AGC
 cys thr leu thr glu arg glu ser his thr ala glu ala gly asp ser gly val val ser
 363/121
 AGC CGA GCT CAA GGT GTC CCG CAC CAC TGT CTC GAA TGC TTT TAA CCG ACC GGA TCC AGA
 ser arg ala gln gly val pro his his cys leu glu cys phe OCH pro thr gly ser arg
 423/141
 AGG AGA AGA TC
 arg arg arg

SEQ ID No.6C

FIGURE 6C

31/11
 CCG TCG GCA ACT TGG CCG CTG AGG TCG GCT TGA TCC CTG GGC CGA GGC GGG TCA GCC AAT
 pro ser ala thr trp pro leu arg ser ala OPA ser leu gly arg gly gly ser ala asn
 61/21
 AGC GGC TCC ATC GGC TTT GCT GGT AGC GGT TCG GCG GGA AGC TAG CGG CGA CGT TGT CGG
 ser gly ser ile gly phe ala gly ser gly ser ala gly ser AMB arg arg arg cys arg
 121/41
 TGG CCG GTG ATA TAT TCG GTC AGA CGG GTA TGG CCG CGG CTG AGG TGA TCT GCG ACA CGC
 trp pro val ile tyr trp val arg arg val trp arg arg leu arg OPA ser ala thr arg
 181/61
 CGC CGC GGT GCT CGA GCC AGG CTT ACG ACC AGG GAA TTT CGA AAA TGT TAT TCA GAA CAT
 arg arg gly ala arg ala arg leu thr thr arg glu phe arg lys cys tyr ser glu his
 241/81
 CTT GTA TCT CTT CTC CGT GCC ACC CCC TAG GTG TAG TGT TTT CGA GTA CCG GCA GAT CCC
 leu val ser leu leu arg ala thr pro AMB val AMB cys phe arg val pro ala asp pro
 301/101
 AGG TTC ACC AGG TCT CAC CAG ATC
 arg phe thr arg ser his gln ile

SEQ ID No.7A

FIGURE 7A

32/11
 CGT CGG CAA CTT GGC CGC TGA GGT CGG CTT GAT CCC TGG GCC GAG GCG GGT CAG CCA ATA
 arg arg gln leu gly arg OPA gly arg leu asp pro trp ala glu ala gly gln pro ile
 62/21
 GCG GCT CCA TCG GCT TTG CTG GTA GCG GTT CGG CGG GAA GCT AGC GGC GAC GTT GTC GGT
 ala ala pro ser ala leu leu val ala val arg arg glu ala ser gly asp val val gly
 122/41
 GGC CGG TGA TAT ATT GGG TCA GAC GGG TAT GGC GGC GGC TGA GGT GAT CTG CGA CAC GCC
 gly arg OPA tyr ile gly ser asp gly tyr gly gly gly OPA gly asp leu arg his ala
 182/61
 GCC GCG GTG CTC GAG CCA GGC TTA CGA CCA GGG AAT TTC GAA AAT GTT ATT CAG AAC ATC
 ala ala val leu glu pro gly leu arg pro gly asn phe glu asn val ile gln asn ile
 242/81
 TTG TAT CTC TTC TCC GTG CCA CCC CCT AGG TGT AGT GTT TTC GAG TAC CGG CAG ATC CCA
 leu tyr leu phe ser val pro pro pro arg cys ser val phe glu tyr arg gln ile pro
 302/101
 GGT TCA CCA GGT CTC ACC AGA TC
 gly ser pro gly leu thr arg

SEQ ID No.7B

FIGURE 7B

33/11
 GTC GGC AAC TTG GCC GCT GAG GTC GGC TTG ATC CCT GGG CCG AGG CGG GTC AGC CAA TAG
 val gly asn leu ala ala glu val gly leu ile pro gly pro arg arg val ser gln AMB
 63/21
 CGG CTC CAT CGG CTT TGC TGG TAG CGG TTC GGC GGC AAG CTA GCG GCG ACG TTG TCG GTG
 arg leu his arg leu cys trp AMB arg phe gly gly lys leu ala ala thr leu ser val
 123/41
 GCC GGT GAT ATA TTG GGT CAG ACG GGT ATG GCG GCG GCT GAG GTG ATC TGC GAC ACG CCG
 ala gly asp ile leu gly gln thr gly met ala ala ala glu val ile cys asp thr pro
 183/61
 CCG CGG TGC TCG AGC CAG GCT TAC GAC CAG GGA ATT TCG AAA ATG TTA TTC AGA ACA TCT
 pro arg cys ser ser gln ala tyr asp gln gly ile ser lys met leu phe arg thr ser
 243/81
 TGT ATC TCT TCT CCG TGC CAC CCC CTA GGT GTA GTG TTT TCG AGT ACC GGC AGA TCC CAG
 cys ile ser ser pro cys his pro leu gly val val phe ser ser thr gly arg ser gln
 303/101
 GTT CAC CAG GTC TCA CCA GAT C
 val his gln val ser pro asp

SEQ ID No.7C

FIGURE 7C

22/185

31/11
 CTT TGC GTG ATG TCC AAT GGC GAA AAC GAC GCC TTG TCA TCG CAA TCG TCA GCA CCG GCC
 leu cys val met ser asn gly glu asn asp ala leu ser ser gln ser ser ala pro ala
 61/21
 TAG TTT TCG CGA TGA CGC TCG TTC TGA CCG GAC TTG TGA ACG GGT TTC GGG TCG AGG CCG
 AMB phe ser arg OPA arg ser phe OPA pro asp leu OPA thr gly phe gly ser arg pro
 121/41
 AGC GAA CCG TCG ATT CCA TGG GTG TCG ACG CAT TCG TGG TCA AGG CCG GCG CGG CAG GAC
 ser glu pro ser ile pro trp val ser thr his ser trp ser arg pro ala arg gln asp
 181/61
 CGT TCC TGG GTT CGA CAC CAT TCG CCC AAA TCG ACC TGC CCC AGG TTG CTC GTG CGC CTG
 arg ser trp val arg his his ser pro lys ser thr cys pro arg leu leu val arg leu
 241/81
 GCG TCT TGG CTG CCG CCC CAC TAG CGA CTG CGC CGT CGA CGA TCC GGC AGG GCA CGT CAG
 ala ser trp leu pro pro his AMB arg leu arg arg arg arg ser gly arg ala arg gln
 301/101
 CGC GAA ACG TCA CCG CGT TCG GGG CAC CAG AGC ACG GAC CCG GCA TGC CGC GGG TCT CGG
 arg glu thr ser pro arg ser gly his gln ser thr asp pro ala cys arg gly ser arg
 361/121
 ACG GTC GGG CGC CAT CGA CGC CGG ACG AGG TCG CGG TGT CGA GCA CGC TGG GCC GAA ACC
 thr val gly arg his arg arg arg thr arg ser arg cys arg ala arg trp ala glu thr
 421/141
 TCG GCG ACG ATC
 ser ala thr ile

SEQ ID No.8A

FIGURE 8A

32/11
 TTT GCG TGA TGT CCA ATG GCG AAA ACG ACG CCT TGT CAT CGC AAT CGT CAG CAC CGG CCT
 phe ala OPA cys pro met ala lys thr thr pro cys his arg asn arg gln his arg pro
 62/21
 AGT TTT CGC GAT GAC GCT CGT TCT GAC CGG ACT TGT GAA CGG GTT TCG GGT CGA GGC CGA
 ser phe arg asp asp ala arg ser asp arg thr cys glu arg val ser gly arg gly arg
 122/41
 GCG AAC CGT CGA TTC CAT GGG TGT CGA CGC ATT CGT GGT CAA GGC CGG CGC GGC AGG ACC
 ala asn arg arg phe his gly cys arg arg ile arg gly gln gly arg arg gly arg thr
 182/61
 GTT CCT GGG TTC GAC ACC ATT CGC CCA AAT CGA CCT GCC CCA GGT TGC TCG TGC GCC TGG
 val pro gly phe asp thr ile arg pro asn arg pro ala pro gly cys ser cys ala trp
 242/81
 CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC GAT CCG GCA GGG CAC GTC AGC
 arg leu gly cys arg pro thr ser asp cys ala val asp asp pro ala gly his val ser
 302/101
 GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CGG ACC CGG CAT GCC GCG GGT CTC GGA
 ala lys arg his arg val arg gly thr arg ala arg thr arg his ala ala gly leu gly
 362/121
 CGG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC GAG CAC GCT GGG CCG AAA CCT
 arg ser gly ala ile asp ala gly arg gly arg gly val glu his ala gly pro lys pro
 422/141
 CGG CGA CGA TC
 arg arg arg

SEQ ID No.8B

FIGURE 8B

REPLACEMENT SHEET (RULE 26)

23/185

33/11
 TTG CGT GAT GTC CAA TGG CGA AAA CGA CGC CTT GTC ATC GCA ATC GTC AGC ACC GGC CTA
 leu arg asp val gln trp arg lys arg arg leu val ile ala ile val ser thr gly leu
 63/21
 GTT TTC GCG ATG ACG CTC GTT CTG ACC GGA CTT GTG AAC GGG TTT CGG GTC GAG GCC GAG
 val phe ala met thr leu val leu thr gly leu val asn gly phe arg val glu ala glu
 123/41
 CGA ACC GTC GAT TCC ATG GGT GTC GAC GCA TTC GTG GTC AAG GCC GGC GCG GCA GGA CCG
 arg thr val asp ser met gly val asp ala phe val val lys ala gly ala ala gly pro
 183/61
 TTC CTG GGT TCG ACA CCA TTC GCC CAA ATC GAC CTG CCC CAG GTT GCT CGT GCG CCT GGC
 phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala arg ala pro gly
 243/81
 GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG ATC CGG CAG GGC ACG TCA GCG
 val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly thr ser ala
 303/101
 CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC GGC ATG CCG CGG GTC TCG GAC
 arg asn val thr ala phe gly ala pro glu his gly pro gly met pro arg val ser asp
 363/121
 GGT CGG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG AGC ACG CTG GGC CGA AAC CTC
 gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu gly arg asn leu
 423/141
 GGC GAC GAT C
 gly asp asp

SEQ ID No.8C

FIGURE 8C

part of the nucleotide sequence of seq8A

1/1
 CAG GTT GCT CGT GCG CCT GGC GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG
 gln val ala arg ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr
 61/21
 ATC CGG CAG GGC ACG TCA GCG CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC
 ile arg gln gly thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro
 121/41
 GGC ATG CCG CGG GTC TCG GAC GGT CGG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG
 gly met pro arg val ser asp gly arg ala pro ser thr pro asp glu val ala val ser
 181/61
 AGC ACG CTG GGC CGA AAC CTC GGC GAC GAT C
 ser thr leu gly arg asn leu gly asp asp

SEQ ID No.8A'

FIGURE 8A'

25/185

sequence Rv2563 predicted by Cole et al. (Nature 393:537-544) and containing seq8A'

```

atg
met
121/41
ctt ttt gcg gct ttg cgt gat gtc caa tgg cga aaa cga cgc ctt gtc atc gca atc gtc
leu phe ala ala leu arg asp val gln trp arg lys arg arg leu val ile ala ile val
181/61
agc acc ggc cta gtt ttc gcg atg acg ctc gtt ctg acc gga ctt gtg aac ggg ttt cgg
ser thr gly leu val phe ala met thr leu val leu thr gly leu val asn gly phe arg
241/81
gtc gag gcc gag cga acc gtc gat tcc atg ggt gtc gac gca ttc gtg gtc aag gcc ggc
val glu ala glu arg thr val asp ser met gly val asp ala phe val val lys ala gly
301/101
gcg gca gga ccg ttc ctg ggt tcg aca cca ttc gcc caa atc gac ctg ccc cag gtt gct
ala ala gly pro phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala
361/121
cgt gcg cct ggc gtc ttg gct gcc gcc cca cta gcg act gcg ccg tcg acg atc cgg cag
arg ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln
421/141
ggc acg tca gcg cga aac gtc acc gcg ttc ggg gca cca gag cac gga ccc ggc atg ccg
gly thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro gly met pro
481/161
cgg gtc tcg gac ggt cgg gcg cca tcg acg ccg gac gag gtc gcg gtg tcg agc acg ctg
arg val ser asp gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu
541/181
ggc cga aac ctc ggc gac gat ctg caa gtg ggt gcg cgc act ttg cgg atc gtc ggc atc
gly arg asn leu gly asp asp leu gln val gly ala arg thr leu arg ile val gly ile
601/201
gtg ccc gag tca acc gcg ctg gca aag att ccc aac atc ttc ctg acc acc gaa ggc cta
val pro glu ser thr ala leu ala lys ile pro asn ile phe leu thr thr glu gly leu
661/221
cag cag ttg gca tac aac gga cag ccg aca atc agt tcg atc ggg atc gac ggg atg ccc
gln gln leu ala tyr asn gly gln pro thr ile ser ser ile gly ile asp gly met pro
721/241
cga cag ctc ccg gac ggc tat cag acc gtc aat cga gcg gat gct gtc agc gat ctg atg
arg gln leu pro asp gly tyr gln thr val asn arg ala asp ala val ser asp leu met
781/261
cgc ccg ttg aag gtc gcg gtg gat gcg atc acg gtt gtg gcg gtc ttg ctg tgg atc gtt
arg pro leu lys val ala val asp ala ile thr val val ala val leu leu trp ile val
841/281
gcg gcg ttg atc gtc ggc tcg gtg gtc tac ctc tct gcg ttg gag ccg ctg cgt gac ttt
ala ala leu ile val gly ser val val tyr leu ser ala leu glu arg leu arg asp phe
901/301
gcg gtg ttc aag gcg atc ggc gtg ccg acg cgc tcg att ctg gcc ggg ctg gcg ctg cag
ala val phe lys ala ile gly val pro thr arg ser ile leu ala gly leu ala leu gln
961/321
gcg gtc gtc gtc gcg ctg ctc gcg gcg gtg gtt ggc ggc atc ctt tcg ctg ctg ttg gcg
ala val val val ala leu leu ala ala val val gly gly ile leu ser leu leu leu ala
1021/341
ccg ttg ttc ccg atg act gtc gtg gta ccc ctg agt gcc ttc gtg gcg cta ccg gcg atc
pro leu phe pro met thr val val val pro leu ser ala phe val ala leu pro ala ile
1081/361
gcg act gtg atc ggt ctg ctg gcc agc gtc gca gga ctg ccg cgc gtg gtg gcg atc gat
ala thr val ile gly leu leu ala ser val ala gly leu arg arg val val ala ile asp
1141/381
ccg gca cta gcg ttc gga ggt ccc tag
pro ala leu ala phe gly gly pro AMB

```

SEQ ID No.8D

FIGURE 8D

REPLACEMENT SHEET (RULE 26)

ORF predicted by Cole et al. (Nature 393:537-544) and containing Rv2563

1/1
tag gtt tca aga agg cct gtg cag gtt tcc gca gcc tgg gcc gcg gcc cca ccg aag agc
AMB val ser arg arg pro val gln val ser ala ala trp ala ala ala pro pro lys ser
61/21
ccg ccg aaa tgg gct aat cgg gtt cgc ttg gct cga tcg ccg atg atc tcg acc gcc acg
pro pro lys trp ala asn arg val arg leu ala arg ser pro met ile ser thr ala thr
121/41
acc gac ccc ctc acc tcg gtc gaa cct cgg cga acc aac gcg gca acg cca gcc cat gat
thr asp pro leu thr ser val glu pro arg arg thr asn ala ala thr pro ala his asp
181/61
cat ttg att ggg tcc acg gaa gca ggt agc ttc cgt cgc atg ctt ttt gcg gct ttg cgt
his leu ile gly ser thr glu ala gly ser phe arg arg met leu phe ala ala leu arg
241/81
gat gtc caa tgg cga aaa cga cgc ctt gtc atc gca atc gtc agc acc ggc cta gtt ttc
asp val gln trp arg lys arg arg leu val ile ala ile val ser thr gly leu val phe
301/101
gcg atg acg ctc gtt ctg acc gga ctt gtg aac ggg ttt cgg gtc gag gcc gag cga acc
ala met thr leu val leu thr gly leu val asn gly phe arg val glu ala glu arg thr
361/121
gtc gat tcc atg ggt gtc gac gca ttc gtg gtc aag gcc ggc gcg gca gga ccg ttc ctg
val asp ser met gly val asp ala phe val val lys ala gly ala ala gly pro phe leu
421/141
ggg tgc aca cca ttc gcc caa atc gac ctg ccc cag gtt gct cgt gcg cct ggc gtc ttg
gly ser thr pro phe ala gln ile asp leu pro gln val ala arg ala pro gly val leu
481/161
gct gcc gcc cca cta gcg act gcg ccg tcg acg atc cgg cag ggc acg tca gcg cga aac
ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly thr ser ala arg asn
541/181
gtc acc gcg ttc ggg gca cca gag cac gga ccc gcc atg ccg ccg gtc tcg gac ggt ccg
val thr ala phe gly ala pro glu his gly pro gly met pro arg val ser asp gly arg
601/201
gcg cca tcg acg ccg gac gag gtc gcg gtg tcg agc acg ctg ggc cga aac ctc ggc gac
ala pro ser thr pro asp glu val ala val ser ser thr leu gly arg asn leu gly asp
661/221
gat ctg caa gtg ggt gcg cgc act ttg ccg atc gtc ggc atc gtg ccc gag tca acc gcg
asp leu gln val gly ala arg thr leu arg ile val gly ile val pro glu ser thr ala
721/241
ctg gca aag att ccc aac atc ttc ctg acc acc gaa ggc cta cag cag ttg gca tac aac
leu ala lys ile pro asn ile phe leu thr thr glu gly leu gln gln leu ala tyr asn
781/261
gga cag ccg aca atc agt tcg atc ggg atc gac ggg atg ccc cga cag ctc ccg gac ggc
gly gln pro thr ile ser ser ile gly ile asp gly met pro arg gln leu pro asp gly
841/281
tat cag acc gtc aat cga gcg gat gct gtc agc gat ctg atg cgc ccg ttg aag gtc gcg
tyr gln thr val asn arg ala asp ala val ser asp leu met arg pro leu lys val ala
901/301
gtg gat gcg atc acg gtt gtg gcg gtc ttg ctg tgg atc gtt gcg gcg ttg atc gtc ggc
val asp ala ile thr val val ala val leu leu trp ile val ala ala leu ile val gly
961/321
tcg gtg gtc tac ctc tct gcg ttg gag ccg ctg cgt gac ttt gcg gtg ttc aag gcg atc
ser val val tyr leu ser ala leu glu arg leu arg asp phe ala val phe lys ala ile
1021/341
ggc gtg ccg acg cgc tcg att ctg gcc ggg ctg gcg ctg cag gcg gtc gtc gtc gcg ctg
gly val pro thr arg ser ile leu ala gly leu ala leu gln ala val val val ala leu
1081/361
ctc gcg gcg gtg gtt ggc ggc atc ctt tcg ctg ctg ttg gcg ccg ttg ttc ccg atg act
leu ala ala val val gly gly ile leu ser leu leu leu ala pro leu phe pro met thr
1141/381
gtc gtg gta ccc ctg agt gcc ttc gtg gcg cta ccg gcg atc gcg act gtg atc ggt ctg
val val val pro leu ser ala phe val ala leu pro ala ile ala thr val ile gly leu
1201/401
ctg gcc agc gtc gca gga ctg ccg cgc gtg gtg gcg atc gat ccg gca cta gcg ttc gga
leu ala ser val ala gly leu arg arg val val ala ile asp pro ala leu ala phe gly
1261/421
ggt ccc tag
gly pro AMB

SEQ ID No.8F

FIGURE 8F

REPLACEMENT SHEET (RULE 26)

sequence of Rv0072 predicted by Cole et al. (Nature 393:537-544) and exhibiting more than 77% similarity with Seq8D'

```

1/1                               31/11
atg ctc ttc gcg gcc ctg cgt gac atg caa tgg aga aag cgc cgc ctg gtc atc acg atc
Met leu phe ala ala leu arg asp met gln trp arg lys arg arg leu val ile thr ile
61/21                               91/31
atc agc acc ggg ctg atc ttc ggg atg acg ctt gtt ttg acc gga ctc gcg aac ggc ttc
ile ser thr gly leu ile phe gly met thr leu val leu thr gly leu ala asn gly phe
121/41                               151/51
cgg gtg gag gcc cgg cac acc gtc gat tcc atg ggt gtc gat gta ttc gtc gtc aga tcc
arg val glu ala arg his thr val asp ser met gly val asp val phe val val arg ser
181/61                               211/71
ggc gct gct gga cct ttt ctg ggt tca ata ccg ttt ccc gat gtt gac ctg gcc cga gtg
gly ala ala gly pro phe leu gly ser ile pro phe pro asp val asp leu ala arg val
241/81                               271/91
gcc gct gaa ccc ggt gtc atg gcc gcg gcc ccg ttg ggc agc gtg ggg acg atc atg aaa
ala ala glu pro gly val met ala ala ala pro leu gly ser val gly thr ile met lys
301/101                               331/111
gaa ggc acg tcg acg cga aac gtc acg gtc ttc ggc gcg ccc gag cac gga cct ggc atg
glu gly thr ser thr arg asn val thr val phe gly ala pro glu his gly pro gly met
361/121                               391/131
cca cgg gtc tca gag ggt cgg tca ccg tcg aaa ccg gac gaa gtc gcg gca tcg agc acg
pro arg val ser glu gly arg ser pro ser lys pro asp glu val ala ala ser ser thr
421/141                               451/151
atg ggc cga cac ctc ggt gac act gtc gag gtc ggc gcg cgc aga ttg cgg gtc gtt ggc
met gly arg his leu gly asp thr val glu val gly ala arg arg leu arg val val gly
481/161                               511/171
att gtg ccg aat tcc acc gcg ctg gcc aag atc ccc aat gtc ttc ctc acg acc gag ggc
ile val pro asn ser thr ala leu ala lys ile pro asn val phe leu thr thr glu gly
541/181                               571/191
tta cag aaa ttg gcg tac aac ggg cag ccg aat atc acg tcc atc ggg atc ata ggt atg
leu gln lys leu ala tyr asn gly gln pro asn ile thr ser ile gly ile ile gly met
601/201                               631/211
ccc cga cag ctg ccg gag ggt tac cag act ttc gat ccg gtg ggc gct gtc aat gat ttg
pro arg gln leu pro glu gly tyr gln thr phe asp arg val gly ala val asn asp leu
661/221                               691/231
gtg cgc cca ttg aag gtc gca gtg aat tcg atc tcg atc gtg gct gtt ttg ctg tgg att
val arg pro leu lys val ala val asn ser ile ser ile val ala val leu leu trp ile
721/241                               751/251
gtg gcg gtg ctg atc gtc ggc tcg gtg gtg tac ctt tcg gct ctt gag cgg cta cgt gac
val ala val leu ile val gly ser val val tyr leu ser ala leu glu arg leu arg asp
781/261                               811/271
ttc gcg gtg ttc aag gcg att ggc acg cca acg cgc tcg att atg gcc ggg ctc gca tta
phe ala val phe lys ala ile gly thr pro thr arg ser ile met ala gly leu ala leu
841/281                               871/291
cag gcg ctg gtc att gcg ttg ctt gcg gcg gtg gtg ggc gtc gtc ctg gcg cag gtg ttg
gln ala leu val ile ala leu leu ala ala val val gly val val leu ala gln val leu
901/301                               931/311
gca cca ctg ttt ccg atg att gtc gcg gta ccc gtc ggt gct tac ctg gcg cta ccg gtg
ala pro leu phe pro met ile val ala val pro val gly ala tyr leu ala leu pro val
961/321                               991/331
gcc gcg atc gtc atc ggt ctg ttc gct agt gtt gcc gga ttg aag cgc gtg gtg acg gtc
ala ala ile val ile gly leu phe ala ser val ala gly leu lys arg val val thr val
1021/341
gat ccc gcg cag gcg ttc gga ggt ccc tag
asp pro ala gln ala phe gly gly pro AMB

```

SEQ ID No.8G

FIGURE 8G

REPLACEMENT SHEET (RULE 26)

Seq8H : ORF predicted by Cole et al. (Nature 393:537-544) and containing seq8G

1/1	31/11
tag cct ctg gga atg ctc ttc gcg gcc ctg cgt gac atg caa tgg aga aag cgc cgc ctg	
AMB pro leu gly met leu phe ala ala leu arg asp met gln trp arg lys arg arg leu	
61/21	91/31
gtc atc acg atc atc agc acc ggg ctg atc ttc ggg atg acg ctt gtt ttg acc gga ctc	
val ile thr ile ile ser thr gly leu ile phe gly met thr leu val leu thr gly leu	
121/41	151/51
gcg aac ggc ttc cgg gtg gag gcc cgg cac acc gtc gat tcc atg ggt gtc gat gta ttc	
ala asn gly phe arg val glu ala arg his thr val asp ser met gly val asp val phe	
181/61	211/71
gtc gtc aga tcc ggc gct gct gga cct ttt ctg ggt tca ata ccg ttt ccc gat gtt gac	
val val arg ser gly ala ala gly pro phe leu gly ser ile pro phe pro asp val asp	
241/81	271/91
ctg gcc cga gtg gcc gct gaa ccc ggt gtc atg gcc gcg gcc ccg ttg ggc agc gtg ggg	
leu ala arg val ala ala glu pro gly val met ala ala ala pro leu gly ser val gly	
301/101	331/111
acg atc atg aaa gaa ggc acg tcg acg cga aac gtc acg gtc ttc ggc gcg ccc gag cac	
thr ile met lys glu gly thr ser thr arg asn val thr val phe gly ala pro glu his	
361/121	391/131
gga cct ggc atg cca cgg gtc tca gag ggt cgg tca ccg tcg aaa ccg gac gaa gtc gcg	
gly pro gly met pro arg val ser glu gly arg ser pro ser lys pro asp glu val ala	
421/141	451/151
gca tcg agc acg atg ggc cga cac ctc ggt gac act gtc gag gtc ggc gcg cgc aga ttg	
ala ser ser thr met gly arg his leu gly asp thr val glu val gly ala arg arg leu	
481/161	511/171
cgg gtc gtt ggc att gtg ccg aat tcc acc gcg ctg gcc aag atc ccc aat gtc ttc ctc	
arg val val gly ile val pro asn ser thr ala leu ala lys ile pro asn val phe leu	
541/181	571/191
acg acc gag ggc tta cag aaa ttg gcg tac aac ggg cag ccg aat atc acg tcc atc ggg	
thr thr glu gly leu gln lys leu ala tyr asn gly gln pro asn ile thr ser ile gly	
601/201	631/211
atc ata ggt atg ccc cga cag ctg ccg gag ggt tac cag act ttc gat ccg gtg ggc gct	
ile ile gly met pro arg gln leu pro glu gly tyr gln thr phe asp arg val gly ala	
661/221	691/231
gtc aat gat ttg gtg cgc cca ttg aag gtc gca gtg aat tcg atc tcg atc gtg gct gtt	
val asn asp leu val arg pro leu lys val ala val asn ser ile ser ile val ala val	
721/241	751/251
ttg ctg tgg att gtg gcg gtg ctg atc gtc ggc tcg gtg gtg tac ctt tcg gct ctt gag	
leu leu trp ile val ala val leu ile val gly ser val val tyr leu ser ala leu glu	
781/261	811/271
cgg cta cgt gac ttc gcg gtg ttc aag gcg att ggc acg cca acg cgc tcg att atg gcc	
arg leu arg asp phe ala val phe lys ala ile gly thr pro thr arg ser ile met ala	
841/281	871/291
ggg ctc gca tta cag gcg ctg gtc att gcg ttg ctt gcg gcg gtg gtg ggc gtc gtc ctg	
gly leu ala leu gln ala leu val ile ala leu leu ala ala val val gly val val leu	
901/301	931/311
gcg cag gtg ttg gca cca ctg ttt ccg atg att gtc gcg gta ccc gtc ggt gct tac ctg	
ala gln val leu ala pro leu phe pro met ile val ala val pro val gly ala tyr leu	
961/321	991/331
gcg cta ccg gtg gcc gcg atc gtc atc ggt ctg ttc gct agt gtt gcc gga ttg aag cgc	
ala leu pro val ala ala ile val ile gly leu phe ala ser val ala gly leu lys arg	
1021/341	1051/351
gtg gtg acg gtc gat ccc gcg cag gcg ttc gga ggt ccc tag	
val val thr val asp pro ala gln ala phe gly gly pro AMB	

SEQ ID No.8H

FIGURE 8H

REPLACEMENT SHEET (RULE 26)

31/11
 CGA GGC CGA GCG AAC CGT CGA TTC CAT GGG TGT CGA CGC ATT CGT GGT CAA GGC CGG CGC
 arg gly arg ala asn arg arg phe his gly cys arg arg ile arg gly gln gly arg arg
 61/21
 GGC AGG ACC GTT CCT GGG TTC GAC ACC ATT CGC CCA AAT CGA CCT GCC CCA GGT TGC TCG
 gly arg thr val pro gly phe asp thr ile arg pro asn arg pro ala pro gly cys ser
 121/41
 TGC GCC TGG CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC GAT CCG GCA GGG
 cys ala trp arg leu gly cys arg pro thr ser asp cys ala val asp asp pro ala gly
 181/61
 CAC GTC AGC GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CGG ACC CGG CAT GCC GCG
 his val ser ala lys arg his arg val arg gly thr arg ala arg thr arg his ala ala
 241/81
 GGT CTC GGA CGG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC GAG CAC GCT GGG
 gly leu gly arg ser gly ala ile asp ala gly arg gly arg gly val glu his ala gly
 301/101
 CCG AAA CCT CGG CGA CGA TC
 pro lys pro arg arg arg

SEQ ID No.9A

FIGURE 9A

32/11
 GAG GCC GAG CGA ACC GTC GAT TCC ATG GGT GTC GAC GCA TTC GTG GTC AAG GCC GGC GCG
 glu ala glu arg thr val asp ser met gly val asp ala phe val val lys ala gly ala
 62/21
 GCA GGA CCG TTC CTG GGT TCG ACA CCA TTC GCC CAA ATC GAC CTG CCC GAG GTT GCT CGT
 ala gly pro phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala arg
 122/41
 GCG CCT GGC GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG ATC CGG CAG GGC
 ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly
 182/61
 ACG TCA GCG CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC GGC ATG CCG CGG
 thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro gly met pro arg
 242/81
 GTC TCG GAC GGT CGG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG ACG ACG CTG GGC
 val ser asp gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu gly
 302/101
 CGA AAC CTC GGC GAC GAT C
 arg asn leu gly asp asp

SEQ ID No.9B

FIGURE 9B

33/11
 AGG CCG AGC GAA CCG TCG ATT CCA TGG GTG TCG ACG CAT TCG TGG TCA AGG CCG GCG CGG
 arg pro ser glu pro ser ile pro trp val ser thr his ser trp ser arg pro ala arg
 63/21
 CAG GAC CGT TCC TGG GTT CGA CAC CAT TCG CCC AAA TCG ACC TGC CCC AGG TTG CTC GTG
 gln asp arg ser trp val arg his his ser pro lys ser thr cys pro arg leu leu val
 123/41
 CGC CTG GCG TCT TGG CTG CCG CCC CAC TAG CGA CTG CGC CGT CGA CGA TCC GGC AGG GCA
 arg leu ala ser trp leu pro pro his AMB arg leu arg arg arg arg ser gly arg ala
 183/61
 CGT CAG CGC GAA ACG TCA CCG CGT TCG GGG CAC CAG AGC ACG GAC CCG GCA TGC CGC GGG
 arg gln arg glu thr ser pro arg ser gly his gln ser thr asp pro ala cys arg gly
 243/81
 TCT CGG ACG GTC GGG CGC CAT CGA CGC CGG ACG AGG TCG CGG TGT CGA GCA CGC TGG GCC
 ser arg thr val gly arg his arg arg arg thr arg ser arg cys arg ala arg trp ala
 303/101
 GAA ACC TCG GCG ACG ATC
 glu thr.ser ala thr ile

SEQ ID No.9C

FIGURE 9C

31/11
 TTA ACG ACT CAG ACG GAA ACG CTT GAA CCG CGA GGT CGC TCC GGA CAC CAA TTT GAC TCG
 leu thr thr gln thr glu thr leu glu pro arg gly arg ser gly his gln phe asp ser
 61/21
 GCT CTT TGG CAA TTG AAG GTG AGC TGC GAG CAG CCG GGT GAC CGC ATC GTT GGC CTT GCC
 ala leu trp gln leu lys val ser cys glu gln pro gly asp arg ile val gly leu ala
 121/41
 ATC AAT CGC CGG CTC GCG GAC GTA GAT AAT CAG CTC ACC GTT GGG ACC GAC CTC GAC CAG
 ile asn arg arg leu ala asp val asp asn gln leu thr val gly thr asp leu asp gln
 181/61
 GGG TCC TTT GTG ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGT CAT CGC CTA AGG
 gly ser phe val thr ala gly leu asp ala asp asp his arg val gly his arg leu arg
 241/81
 CTA CCG TTC TGA CCT GGG GCT GCG TGG GCG CCG ACG ACG TGA GGC ACG TCA TGT CTC AGC
 leu pro phe OPA pro gly ala ala trp ala pro thr thr OPA gly thr ser cys leu ser
 301/101
 GGC CCA CCG CCA CCT CGG TCG CCG GCA GTA TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC
 gly pro pro pro pro arg ser pro ala val cys gln his val gln met thr pro arg ser
 361/121
 CTT GTT CGC ATC GTT GGT GTC GTG GTT GCG ACG ACC TTG GCG CTG GTG AGC GCA CCC GCC
 leu val arg ile val gly val val val ala thr thr leu ala leu val ser ala pro ala
 421/141
 GGC GGT CGT GCC GCG CAT GCG GAT C
 gly gly arg ala ala his ala asp

SEQ ID No.10A

FIGURE 10A

31/185

32/11
TAA CGA CTC AGA CGG AAA CGC TTG AAC CGC GAG GTC GCT CCG GAC ACC AAT TTG ACT CGG
OCH arg leu arg arg lys arg leu asn arg glu val ala pro asp thr asn leu thr arg
62/21 92/31
CTC TTT GGC AAT TGA AGG TGA GCT GCG AGC AGC CGG GTG ACC GCA TCG TTG GCC TTG CCA
leu phe gly asn OPA arg OPA ala ala ser ser arg val thr ala ser leu ala leu pro
122/41 152/51
TCA ATC GCC GGC TCG CGG ACG TAG ATA ATC AGC TCA CCG TTG GGA CCG ACC TCG ACC AGG
ser ile ala gly ser arg thr AMB ile ile ser ser pro leu gly pro thr ser thr arg
182/61 212/71
GGT CCT TTG TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GTC ATC GCC TAA GGC
gly pro leu OPA leu pro gly leu thr arg thr thr glu ser val ile ala OCH gly
242/81 272/91
TAC CGT TCT GAC CTG GGG CTG CGT GGG CGC CGA CGA CGT GAG GCA CGT CAT GTC TCA GCG
tyr arg ser asp leu gly leu arg gly arg arg arg arg glu ala arg his val ser ala
302/101 332/111
GCC CAC CGC GAC CTC GGT CGC CGG CAG TAT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCC
ala his arg his leu gly arg arg gln tyr val ser met cys arg OPA leu his ala ala
362/121 392/131
TTG TTC GCA TCG TTG GTG TCG TGG TTG CGA CGA CCT TGG CGC TGG TGA GCG CAC CCG CCG
leu phe ala ser leu val ser trp leu arg arg pro trp arg trp OPA ala his pro pro
422/141
GCG GTC GTG CCG CGC ATG CGG ATC
ala val val pro arg met arg ile

SEQ ID No.10B

FIGURE 10B

33/11
AAC GAC TCA GAC GGA AAC GCT TGA ACC GCG AGG TCG CTC CGG ACA CCA ATT TGA CTC GGC
asn asp ser asp gly asn ala OPA thr ala arg ser leu arg thr pro ile OPA leu gly
63/21 93/31
TCT TTG GCA ATT GAA GGT GAG CTG CGA GCA GCC GGG TGA CCG CAT CGT TGG CCT TGC CAT
ser leu ala ile glu gly glu leu arg ala ala gly OPA pro his arg trp pro cys his
123/41 153/51
CAA TCG CCG GCT CGC GGA CGT AGA TAA TCA GCT CAC CGT TGG GAC CGA CCT CGA CCA GGG
gln ser pro ala arg gly arg arg OCH ser ala his arg trp asp arg pro arg pro gly
183/61 213/71
GTC CTT TGT GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT CGG TCA TCG CCT AAG GCT
val leu cys asp cys arg ala OPA arg gly arg pro gln ser arg ser ser pro lys ala
243/81 273/91
ACC GTT CTG ACC TGG GGC TGC GTG GGC GCC GAC GAC GTG AGG CAC GTC ATG TCT CAG CGG
thr val leu thr trp gly cys val gly ala asp asp val arg his val met ser gln arg
303/101 333/111
CCC ACC GCC ACC TCG GTC GCC GGC AGT ATG TCA GCA TGT GCA GAT GAC TCC ACG CAG CCT
pro thr ala thr ser val ala gly ser met ser ala cys ala asp asp ser thr gln pro
363/121 393/131
TGT TCG CAT CGT TGG TGT CGT GGT TGC GAC GAC CTT GGC GCT GGT GAG CGC ACC CGC CGG
cys ser his arg trp cys arg gly cys asp asp leu gly ala gly glu arg thr arg arg
423/141
CGG TCG TGC CGC GCA TGC GGA TC
arg ser cys arg ala cys gly

SEQ ID No.10C

FIGURE 10C

32/185

31/11
 CCC GAA GAG GTC CCC CGT TTT GTT AAT TTT TAA AAA ATT TGT GTC ACA AAC CGG GGT ACC
 pro glu glu val pro arg phe val asp phe OCH lys ile cys val thr lys arg gly thr
 61/21 91/31
 AAG GCA TAA AAC CTA GTA CCT GGG GCG GCG GAT TCA ACG AAA ACC GAG TGG GGG TAG TCA
 lys ala OCH asn leu val pro gly ala ala asp ser thr lys thr glu trp gly AMB ser
 121/41 151/51
 GGG GCG TGC ATT CCG ACG ACC CTG TAC GAC CCG CTG GTG GCA ACG CCG ATG AGT GCG CCG
 gly ala cys ile pro thr thr leu tyr asp pro leu val ala thr pro met ser ala pro
 181/61 211/71
 ACG AAG GCC GAG CGA CGG GCT GCC GGC GCT GAC CGC CGC GGA AGC CGC CGA GTG CAT GGT
 thr lys ala glu arg arg ala ala gly ala asp arg arg gly ser arg arg val asp gly
 241/81 271/91
 CAC CAC CGC CCG CAC CCG ACC GGT ACG GAT CGC GCC TCG GGT TAC CGT CGC CGT CAA CGC
 his his arg pro his pro thr gly thr asp arg ala ser gly tyr arg arg arg gln arg
 301/101 331/111
 GCT GGA CAG CAT CGG TCC CCG CTG GGT CAA TGC ACT CAT GCA GCG CCG CAA CGA ACA GCT
 ala gly gln his arg ser pro leu gly gln cys thr his ala ala pro gln arg thr ala
 361/121 391/131
 CAA CCC TTG AAC CGG GTC CCG GCC TGC CGA CCC TCG GCC GCC GGC GTG CCG CTA CGT GAT
 gln pro leu asn arg val pro ala cys arg pro ser ala ala gly val pro leu arg asp
 421/141 451/151
 AGA CAC AGG GCC ATG GAA ATC CTG GCC AGC CGG ATG CTA CTT CGG CCG GCG GAC TAT CAG
 arg his arg ala met glu ile leu ala ser arg met leu leu arg pro ala asp tyr gln
 481/161
 CGG TCG CTG AGC TTC TAC CGT GAC CAG ATC
 arg ser leu ser phe tyr arg asp gln ile

SEQ ID No.11A

FIGURE 11A

32/11
 CCG AAG AGG TCC CCC GTT TTG TTA ATT TTT AAA AAA TTT GTG TCA CAA AGC GGG GTA CCA
 pro lys arg ser pro val leu leu ile phe lys lys phe val ser gln ser gly val pro
 62/21 92/31
 AGG CAT AAA ACC TAG TAC CTG GGG CGG CGG ATT CAA CGA AAA CCG AGT GGG GGT AGT CAG
 arg his lys thr AMB tyr leu gly arg arg ile gln arg lys pro ser gly gly ser gln
 122/41 152/51
 GGG CGT GCA TTC CGA CGA CCC TGT ACG ACC CGC TGG TGG CAA CGC CGA TGA GTG CCC CGA
 gly arg ala phe arg arg pro cys thr thr arg trp trp gin arg arg OPA val arg arg
 182/61 212/71
 CGA AGG CCG AGC GAC GGG CTG CCG GCG CTG ACC GCC GCG GAA GCC GCC GAG TGG ATG GTC
 arg arg pro ser asp gly leu pro ala leu thr ala ala glu ala ala glu trp met val
 242/81 272/91
 ACC ACC GCC CGC ACC CGA CCG GTA CGG ATC CGC CCT CGG GTT ACC GTC GCC GTC ACC GCG
 thr thr ala arg thr arg pro val arg ile ala pro arg val thr val ala val asn ala
 302/101 332/111
 CTG GAC AGC ATC GGT CCC CGC TGG GTC AAT GCA CTC ATG CAG CGC CGC AAC GAA CAG CTC
 leu asp ser ile gly pro arg trp val asn ala leu met gln arg arg asn glu gln leu
 362/121 392/131
 AAC CCT TGA ACC GGG TCC CGG CCT GCC GAC CCT CGG CCG CCG GCG TGC CGC TAC GTG ATA
 asn pro OPA thr gly ser arg pro ala asp pro arg pro pro ala cys arg tyr val ile
 422/141 452/151
 GAC ACA GGG CCA TGG AAA TCC TGG CCA GCC GGA TGC TAC TTC GCC CGG CGG ACT ATC AGC
 asp thr gly pro trp lys ser trp pro ala gly cys tyr phe gly arg arg thr ile ser
 482/161
 GGT CGC TGA GCT TCT ACC GTG ACC AGA TC
 gly arg OPA ala ser thr val thr arg

SEQ ID No.11B

FIGURE 11B

33/11
 CGA AGA GGT CCC CCG TTT TGT TAA TTT TTA AAA AAT TTG TGT CAC AAA GCG GGG TAC CAA
 arg arg gly pro pro phe cys OCH phe leu lys asn leu cys his lys ala gly tyr gln
 63/21 93/31
 GGC ATA AAA CCT AGT ACC TGG GGC GGC GGA TTC AAC GAA AAC CGA GTG GGG GTA GTC AGG
 gly ile lys pro ser thr trp gly gly gly phe asn glu asn arg val gly val val arg
 123/41 153/51
 GGC GTG CAT TCC GAC GAC CCT GTA CGA CCC GCT GGT GGC AAC GCC GAT GAG TGC GCC GAC
 gly val his ser asp asp pro val arg pro ala gly gly asn ala asp glu cys ala asp
 183/61 213/71
 GAA GGC CGA GCG ACG GGC TGC CGG CGC TGA CCG CCG CGG AAG CCG CCG AGT GGA TGG TCA
 glu gly arg ala thr gly cys arg arg OPA pro pro arg lys pro pro ser gly trp ser
 243/81 273/91
 CCA CCG CCC GCA CCC GAC CGG TAC GGA TCG CGC CTC GGG TTA CCG TCG CCG TCA ACG CGC
 pro pro pro ala pro asp arg tyr gly ser arg leu gly leu pro ser pro ser thr arg
 303/101 333/111
 TGG ACA GCA TCG GTC CCC GCT GGG TCA ATG CAC TCA TGC AGC GCC GCA ACG AAC AGC TCA
 trp thr ala ser val pro ala gly ser met his ser cys ser ala ala thr asn ser ser
 363/121 393/131
 ACC CTT GAA CCG GGT CCC GGC CTG CCG ACC CTC GGC CGC CGG CGT GCC GCT ACG TGA TAG
 thr leu glu pro gly pro gly leu pro thr leu gly arg arg arg ala ala thr OPA AMB
 423/141 453/151
 ACA CAG GGC CAT GGA AAT CCT GGC CAG CCG GAT GCT ACT TCG GCC GGC GGA CTA TCA GCG
 thr gln gly his gly asn pro gly gln pro asp ala thr ser ala gly gly leu ser ala
 483/161
 GTC GCT GAG CTT CTA CCG TGA CCA GAT C
 val ala glu leu leu pro OPA pro asp

SEQ ID No.11C

FIGURE 11C

part of the nucleotide sequence of Seq11

1/1 31/11
 CGT CGC CGT CAA CGC GCT GGA CAG CAT CGG TCC CCG CTG GGT CAA TGC ACT CAT GCA GCG
 arg arg arg gln arg ala gly gln his arg ser pro leu gly gln cys thr his ala ala
 61/21 91/31
 CCG CAA CGA ACA GCT CAA CCC TTG AAC CGG GTC CCG GCC TGC CGA CCC TCG GCC GCC GGC
 pro gln arg thr ala gln pro leu asn arg val pro ala cys arg pro ser ala ala gly
 121/41 151/51
 GTG CCG CTA CGT GAT AGA CAC AGG GCC ATG GAA ATC CTG GCC AGC CGG ATG CTA CTT CGG
 val pro leu arg asp arg his arg ala met glu ile leu ala ser arg met leu leu arg
 181/61 211/71
 CCG GCG GAC TAT CAG CGG TCG CTG AGC TTC TAC CGT GAC CAG ATC
 pro ala asp tyr gln arg ser leu ser phe tyr arg asp gln ile

SEQ ID No.11A'

FIGURE 11A'

1/1	31/11
GTC GCC GTC AAC GCG CTG GAC AGC ATC GGT	CCC CGC TGG GTC AAT GCA CTC ATG CAG CGC
val ala val asn ala leu asp ser ile gly	pro arg trp val asn ala leu met gln arg
61/21	91/31
CGC AAC GAA CAG CTC AAC CCT TGA ACC GGG	TCC CGG CCT GCC GAC CCT CGG CCG CCG GCG
arg asn glu gln leu asn pro OPA thr gly	ser arg pro ala asp pro arg pro pro ala
121/41	151/51
TGC CGC TAC GTG ATA GAC ACA GGG CCA TGG	AAA TCC TGG CCA GCC GGA TGC TAC TTC GGC
cys arg tyr val ile asp thr gly pro trp	lys ser trp pro ala gly cys tyr phe gly
181/61	211/71
CGG CGG ACT ATC AGC GGT CGC TGA GCT TCT	ACC GTG ACC AGA TC
arg arg thr ile ser gly arg OPA ala ser	thr val thr arg

SEQ ID No.11B'

FIGURE 11B'

1/1	31/11
TCG CCG TCA ACG CGC TGG ACA GCA TCG GTC	CCC GCT GGG TCA ATG CAC TCA TGC AGC GCC
ser pro ser thr arg trp thr ala ser val	pro ala gly ser met his ser cys ser ala
61/21	91/31
GCA ACG AAC AGC TCA ACC CTT GAA CCG GGT	CCC GGC CTG CCG ACC CTC GGC CGC CGG CGT
ala thr asn ser ser thr leu glu pro gly	pro gly leu pro thr leu gly arg arg arg
121/41	151/51
GCC GCT ACG TGA TAG ACA CAG GGC CAT GGA	AAT CCT GGC CAG CCG GAT GCT ACT TCG GCC
ala ala thr OPA AMB thr gln gly his gly	asn pro gly gln pro asp ala thr ser ala
181/61	211/71
GGC GGA CTA TCA GCG GTC GCT GAG CTT CTA	CCG TGA CCA GAT C
gly gly leu ser ala val ala glu leu leu	pro OPA pro asp

SEQ ID No.11C'

FIGURE 11C'

sequence Rv0546c predicted by Cole et al. (Nature 393:537-544) and containing Seq11A'

1/1	31/11
atg gaa atc ctg gcc agc cgg atg cta ctt	cgg ccg gcg gac tat cag cgg tcg ctg agc
Met glu ile leu ala ser arg met leu leu	arg pro ala asp tyr gln arg ser leu ser
61/21	91/31
ttc tac cgt gac cag atc ggg ctg gcg att	gcc cgt gaa tac ggg gcc gcc aca gtg ttt
phe tyr arg asp gln ile gly leu ala ile	ala arg glu tyr gly ala gly thr val phe
121/41	151/51
ttc gcc ggt cag tca ctg ctc gaa ctg gcc	ggt tac gcc gag ccg gac cat tcg ccg gga
phe ala gly gln ser leu leu glu leu ala	gly tyr gly glu pro asp his ser arg gly
181/61	211/71
cct ttt ccc gcc gcg ctg tgg ctg cag gtg	cgc gac ctc gag gct acc cag acc gag ctg
pro phe pro gly ala leu trp leu gln val	arg asp leu glu ala thr gln thr glu leu
241/81	271/91
gtc agc cga gcc gtg tcg atc gct cgc gag	ccc cgc cgc gaa ccg tgg gcc ctg cac gag
val ser arg gly val ser ile ala arg glu	pro arg arg glu pro trp gly leu his glu
301/101	331/111
atg cat gtg acc gac cca gac ggg atc aca	ctg ata ttc gtc gag gtt ccc gag ggt cac
met his val thr asp pro asp gly ile thr	leu ile phe val glu val pro glu gly his
361/121	
ccg ctg cgt aca gac acc cgg gcg tga	
pro leu arg thr asp thr arg ala OPA	

SEQ ID No.11D

FIGURE 11D